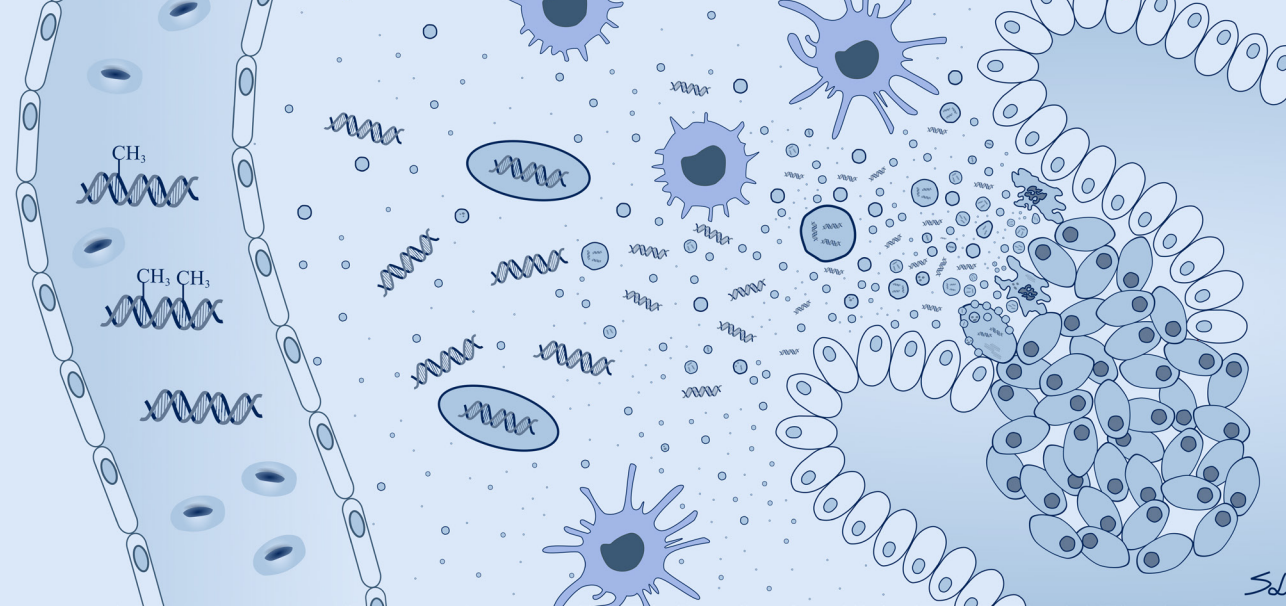
Diagnostic and Prognostic Potential of Hypermethylated DNA in Colorectal Cancer

Rasmussen, Simon Ladefoged

DOI (link to publication from Publisher):
10.5278/vbn.phd.med.00104

Publication date:
2017

Document Version
Publisher's PDF, also known as Version of record



DIAGNOSTIC AND PROGNOSTIC POTENTIAL OF HYPERMETHYLATED DNA IN COLORECTAL CANCER

A PLASMA BASED BIOMARKER STUDY

BY
SIMON LADEFOGED RASMUSSEN

DISSERTATION SUBMITTED 2017

AALBORG UNIVERSITY
DENMARK

DIAGNOSTIC AND PROGNOSTIC POTENTIAL OF HYPERMETHYLATED DNA IN COLORECTAL CANCER

A PLASMA BASED BIOMARKER STUDY

by

Simon Ladefoged Rasmussen

AALBORG UNIVERSITY
DENMARK

Dissertation submitted October 2017

Dissertation submitted: November 2017

PhD supervisor: Prof. M.D. DMSc. Ole Thorlacius-Ussing
Dept. Gastrointestinal Surgery
Aalborg University Hospital

Assistant PhD supervisors: M.D. Ph.D. Henrik Bygum Krarup
Dept. Molecular Diagnostics
Aalborg University Hospital

M.Sc. Poul Henning Madsen
Dept. Molecular Diagnostics
Aalborg University Hospital

M.D. Ph.D. Kåre Gotschalck Sunesen
Dept. Gastrointestinal Surgery
Aalborg University Hospital

PhD committee: Ursula Falkmer, Professor, MD, PhD
Aalborg University Hospital
Aalborg University

Per J. Nilsson, Senior Lecturer, Associate Professor of Surgery, MD, PhD
Karolinska Institutet, Stockholm

Mads Thomassen, Associate Professor, PhD
Odense University Hospital
University of Southern Denmark

PhD Series: Faculty of Medicine, Aalborg University
Department: Department of Clinical Medicine

ISSN (online): 2246-1302
ISBN (online): 978-87-7210-100-2

Published by:
Aalborg University Press
Skjernvej 4A, 2nd floor
DK – 9220 Aalborg Ø
Phone: +45 99407140
aauf@forlag.aau.dk
forlag.aau.dk

Printed in Denmark by Rosendahls, 2017

For Line

PREFACE

The current PhD dissertation could very well be a chance encounter. It was conceived during a summer holiday in 2012. During the last week of my clinical rotation at The Department of Gastrointestinal Surgery, Professor Ole Thorlacius-Ussing presented an idea surrounding circulating DNA as biomarkers for colorectal cancer. I felt very privileged since I was chosen for this project. However, I later realised, that I was merely the second choice… Nonetheless, my efforts were not in vain, and the pilot study eventually led to my enrolment as a Ph.D. fellow at Aalborg University.

First, I would like to thank my main supervisor Professor Ole Thorlacius-Ussing, for giving me the opportunity and introducing me to the world of clinical research. Your knowledge, work ethics, and constant support is an inspiration. You have opened your house to me, in times of need, and introduced me to your lovely wife Lise - I always feel at home in your company. I feel very fortunate, to call you my mentor and my friend.

Second, I very much appreciate the efforts made by Mogens Stender. Your meticulous characterisation of the study cohort is admirable, and I cannot thank you enough for letting me use your precious work to conduct my thesis.

Third, I am exceedingly grateful for my collaboration with Inge Søkilde Pedersen, Poul Henning Madsen, and Henrik Bygum Krarup. You have helped me understand fundamental molecular methods and their application and pushed me to increase my knowledge surrounding diagnostic tests. I hope the future work will provide as much insight as the former years have given.

Fourth, I have been blessed with the best colleagues at The Research Unit at the Department of Gastrointestinal surgery. I extend my complete and sincere thanks to: Stine Dam Henriksen, for being my sparring partner and hypermethylation buddy, Karina Frahm Kirk; for simply being brilliant. Henriette Strøm Kahr, for your keen eye to detail and likeminded epidemiological interest. Lone Schmidt Sørensen, for your joyful spirit and endless efforts to ensure my interest in colorectal surgery. Sabrina Just Kousgaard, for your delightful enthusiasm. Anni Bahnsen, for your caring personality and moral support. Ann Hauberg, for your helping nature. June Lundtoft, for your constant and meticulous assistance. Martin Berg Johansen, for you apt statistical mind and patience in teaching the unenlightened. Kåre Gotschalck Sunesen, for your intelligence, enormous knowledge, and continuous effort in keeping me on track. David Straarup, for your positive nature. Anders Christian Larsen, for being a close support in the initial stages of the project.

Finally, I would like to extend my whole-hearted appreciation and gratitude to my parents Karl Aage Rasmussen and Birte Ladefoged Rasmussen. Your endless care and consideration cannot be appreciated enough.

Simon Ladefoged Rasmussen 2017

ENGLISH SUMMARY

Globally, colorectal cancer (CRC) is the third most common cancer with an incidence of approximately 1.3 million cases annually. In 2014, screening for CRC was implemented in Denmark, and during the same year, 5,186 people were diagnosed with CRC. The survival of CRC patients is closely related to the stage at the time of diagnosis prompting the importance of early detection. CRC takes several years to develop, and have known precursor lesions (adenomas). Like most solid cancers, CRC develops through several molecular alterations. Known driver mutations include *APC* and *KRAS* mutation, which provide growth advantages for neoplastic cells fuelling the adenoma to carcinoma sequence.

DNA hypermethylation is a normal molecular phenomenon in eukaryotic cells. Under normal conditions, this phenomenon is under strict regulation, and primarily involved in regulating embryological development. In different disease states, however, hypermethylation of the DNA promoter region infers a transcriptional down-regulation, and hypermethylation of tumour suppressor genes is a frequent event in the early stages of CRC development. DNA fragments are constantly present in the circulation, termed cell-free DNA. The cell-free DNA can be analysed for aberrant methylation which could be used as biomarkers for CRC detection, prognostication, and disease progression. The aim of this thesis was to ascertain already proven CRC hypermethylation biomarkers with regards to their capabilities for CRC detection and prognostication.

In order to select biomarkers for further analysis we conducted a systematic literature review. Previous studies have identified approximately 70 different hypermethylated genes promoter regions as blood or stool based biomarkers for CRC. Some of these individual biomarkers seems to be able to distinguish early stage CRC patients from healthy controls (e.g. *APC*, *NEUROG1*, *RASSF1A*, *RASSF2A*, *SDC2*, *SEPT9*, and *THBD*). *SEPT9* is already commercialised as a blood based biomarker for CRC detection. However, a large-scale cross-sectional study in 1.457 screening subjects revealed a sensitivity of only 48.2%. Combined hypermethylation biomarker assays have revealed sensitivities and specificities above 90% (*APC, MGMT, RASSF2A, and WIF1*) suggesting that the combination of multiple hypermethylation biomarkers could increase clinical utility.

We evaluated 30 of the DNA promoter regions found in the aforementioned literature study. Through the analysis of plasma samples from 193 CRC patients and 102 colonoscopy verified healthy controls, we found that each individual cell-free hypermethylated biomarker only provided a sensitivity of 30% at a reasonable specificity. Seven hypermethylated promoter regions (*ALX4*, *BMP3*, *NPTX2*, *RARB*, *SDC2*, *SEPT9*, and *VIM*), however, had the ability to detect CRC with 90.7% sensitivity and

72.5% specificity. The panel had a similar performance when analysing only stage I and II CRC patients, indicating that these seven biomarkers could be used for CRC screening. Whether this result is reproducible is pending a subsequent validation study.

Moreover, we found, that the majority of the hypermethylated promoter regions inferred a decreased survival compared with patients not having these biomarkers. The five-year survival was 79.7% (63.8; 82.3) in patients with a low number of circulating hypermethylated promoter regions (< 5) compared with 40.6% (23.8; 56.8) in patients with a high number (> 10). Two of the DNA biomarkers (*RARB* and *RASSF1A*) were significant predictors of poor overall survival regardless of CRC stage at the time of diagnosis (hazard ratios of 1.99 [1.07; 3.72] and 3.35 [1.76; 6.38] respectively) suggesting aggressive disease. A high number of hypermethylated promoter regions was also associated with an increased risk of CRC recurrence (both metastatic and regional) with a stage-adjusted three-year risk ratio of 1.91 (0.91; 4.04). This difference was, however, not statistically significant.

In conclusion, individual cell-free hypermethylated DNA promoter regions have a limited ability to detect CRC but a panel of seven hypermethylated promoter regions show promise as a blood-based biomarker for CRC detection. Moreover, the number of circulating hypermethylated promoter regions was associated with a decreased overall survival, and a trend towards an increased risk of disease recurrence. These findings are in line with previous findings and demand further validation to ensure the clinical utility of the biomarkers in CRC detection and prognostication.

DANISH SUMMARY

Kræft i tyk- og endetarmen (KRC) er en af hyppigste og mest dødelige kræftformer med op imod 1.3 millioner tilfælde årligt på verdensplan. I Danmark er der derfor indført screening for KRC, og i 2014 fik 5,186 danskere konstateret sygdommen. Overlevelsen er tæt korreleret med stadiet på diagnosetidspunktet, hvilket nødvendiggør tidlig og sikker diagnostik. KRC udvikler sig over en flerårig periode og har kendte forstadier (adenomer og polypper). Ved transformationen fra normal celle til kræftcelle opstår der en række molekylære ændringer i cellernes arvemateriale (DNA). Mutationer i generne *APC* og *KRAS* ses ofte i KRC og giver kræftcellerne en vækstfordel i forhold til de omkringliggende normale celler. Disse mutationer er drivkræfterne for omdannelse af adenomer til karcinomer (KRC).

DNA hypermethylering er en normal reguleringsmekanisme i eukaryote celler. Det ses som en normal del af den humane embryologiske udvikling, og det medvirker til korrekt regulering af, hvilke gener der udtrykkes. Hypermethylering af startsekvensen af forskellige gener fører til inaktivering. I den tidlige udvikling af forskellige kræftformer ses det, at gener, som er involveret i celledeling og differentiering, bliver hypermethylerede og dermed gjort inaktive. Når disse gener bliver inaktive, vil cellen ikke længere være underlagt de sædvanlige kontrolmekanismer, og kan derfor potentielt udvikle sig til en kræftcelle. Det er muligt at isolere humant DNA fra blodprøver. Dette DNA kan derpå analyseres for hypermethylering, hvilket potentielt kan bruges som markør for KRC. Projekts formål var, at undersøge om DNA hypermethylering målt i blodprøver kunne bruges til diagnosticere tyk- og endetarmskræft. Herudover var det hensigten, at undersøge om disse blodbårne markører kunne anvendes til at differentiere mellem KRC patienter med en god prognose og patienter med en dårlig prognose.

De individuelle KRC specifikke hypermethyleringsmarkører blev udvalgt på baggrund af et litteraturstudie. Tidligere studier har identificeret ca. 70 forskellige hypermethylerede gener som potentielle KRC markører i blod- og afføringsprøver. Nogle af disse individuelle markører er fundet i blodet i tidlige stadier af KRC (f.eks. *APC*, *NEUROG1*, *RASSF1A*, *RASSF2A*, *SDC2*, *SEPT9*, and *THBD*), og *SEPT9* anvendes allerede i en kommerciel test. Hvorvidt *SEPT9* er en klinisk anvendelig markør er dog til debat. Et stort tværsnitsstudie (1.457 screeningspatienter) viste at *SEPT9* kun havde en sensitivitet på 48.2%. Andre studier, som har kombineret flere markører (*APC*, *MGMT*, *RASSF2A*, og *WIF1*), har vist mere lovende resultater med sensitivitets og specificitetsestimater over 90%. Dette tyder på, at en kombination af flere hypermethyleringsmarkører bør anvendes, for at højne sensitiviteten.

Fra vores litteraturstudie udvalgte vi 30 markører til videre analyse. Disse markører undersøgte vi derpå i blodprøver fra 193 KRC patienter samt 102 kontrolpersoner.

Ud fra disse blodprøver fandt vi, at hver enkelt markør kun havde en sensitivitet på ca. 30%. Dog kunne syv af de analyserede markører (*ALX4*, *BMP3*, *NPTX2*, *RARB*, *SDC2*, *SEPT9*, and *VIM*) adskille de to grupper med 90.7% sensitivitet og 72.5% specificitet. Markørerne kunne identificere KRC patienter i tidligt stadie (stadie I og II) med lignende performance, hvilket indikerer at markørerne kan bruges til screeningen for KRC. Om dette markørpanel kan anvendes i det danske screeningsprogram afventer et opfølgningsstudie.

Ydermere fandt vi, at langt hovedparten af de undersøgte hypermethyleringsmarkører var udtryk for en dårlig prognose. Femårsoverlevelsen hos KRC patienter med under fem markører i blodet var 79.2% (63.8; 82.3) sammenlignet med 40.6% (23.8; 56.8) hos patienter med mere end ti markører. Specielt var tilstedeværelsen af *RARB* og *RASSF1A* hypermethylering et dårligt prognostisk tegn (stadiejusterede hazard rater på hhv. 1.99, 95%CI [1.07; 3.72] og 3.35, 95%CI [1.76; 6.38]). Et højt antal hypermethyleringer var endvidere et udtryk for øget for recidivrisiko med en stadiejusteret tre års risiko ratio på 1.91 (0.91; 4.04), hvilket dog ikke var statistisk signifikant.

Som konklusion kan individuelle hypermethyleringsmarkører ikke anvendes som blodbårne markører for KRC. Det ser dog ud til, at et panel af syv markører kan anvendes som screeningstest for KRC. Ydermere er et højt antal af disse markører et udtryk for dårlig prognose og en øget risiko for recidiv. Disse fund er sammenlignelige med tidligere studier og fordrer validering i fremtidige studier for at sikre deres kliniske brugbarhed.

LIST OF STUDIES

I. Hypermethylated DNA, a Diagnostic and Prognostic Marker for Colorectal Cancer – a systematic review.
S. L. Rasmussen, H. B. Krarup, K. G. Sunesen, I. S. Pedersen, P. H. Madsen, O. Thorlacius-Ussing.
Colorectal Disease. 2016 Jun;18(6):549-61.
DOI: 10.1111/codi.13336

II. Hypermethylated DNA, a Circulating Biomarker for Colorectal Cancer Detection.
S. L. Rasmussen, H. B. Krarup, K. G. Sunesen, M. B. Johansen, M. T. Stender, I. S. Pedersen, P. H. Madsen, O. Thorlacius-Ussing.
PLoS One. 2017 Jul 10;12(7):e0180809
DOI: 10.1371/journal.pone.0180809

III. The Prognostic Utility of Circulating Hypermethylated DNA in Colorectal Cancer.
S. L. Rasmussen, H. B. Krarup, K. G. Sunesen, M. B. Johansen, M. T. Stender, I. S. Pedersen, P. H. Madsen, O. Thorlacius-Ussing
Submitted for publication in *Oncotarget*.

TABLE OF CONTENTS

ABBREVIATIONS

95%CI	95% confidence interval
AJCC	American joint committee on cancer
AUC	Area under the receiver operating characteristics curve
CEA	Carcinoembryonic antigen
CIMP	CpG island methylator phenotype
CMS	Consensus molecular subtypes
CpG	Cytosine phosphate guanine
CRC	Colorectal cancer
CT	Computed tomography
Ct	Cycle threshold
DNMT	DNA methyltransferase
EGFR	Epidermal growth factor receptor
EWAS	Epigenome wide association studies
FAP	Familial adenomatous polyposis
GI	Gastrointestinal tract
GWAS	Genome wide associations studies
HNPCC	Hereditary nonpolyposis colorectal cancer
HR	Hazard ratio
iFOBT	Immunochemical faecal occult blood test
MMR	Mismatch repair
MRI	Magnetic resonance imaging
MSI	Microsatellite instability
OR	Odds ratio
PCR	Polymerase chain reaction
PET/CT	Positron emission tomography computed tomography
RFA	Radio frequency ablation
ROC	Receiver operating characteristics curve
RR	Risk ratio
SNP	Single nucleotide polymorphism

TABLES AND FIGURES

CHAPTER 1. INTRODUCTION

1.1. COLORECTAL CANCER

Colorectal cancer (CRC) arises from the cell lining of the large intestine. The current thesis focuses on the epigenetic alterations in CRC, and the following sections will provide the reader with an overview of the different aspects of the disease.

1.1.1. ANATOMY OF THE LARGE INTESTINE

The human gastrointestinal tract (GI) is a hollow organ, arising in the oral cavity and ending in the anal canal. The primary function is to ensure transportation, digestion, and absorption of ingested nutrients. Moreover, it has a wide range of secondary functions encompassing the immune, endocrine, and cardiovascular system.

The GI tract is divided into several sections, due to embryological origin, anatomical location, physical appearance, and physiological function. The large intestine is the last part of the GI tract. It extends from the terminal part of the small intestine (the ileum) to the anus. It is approximately 1.5 m in length, and it is relatively more fixated than the small bowel.

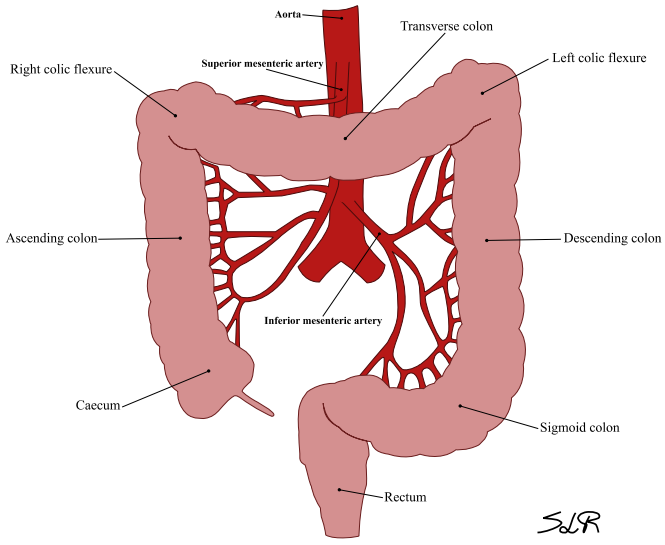


Figure 1 Schematic drawing of the colon and rectum, with blood-supply

The primitive intestinal tube develops from the endodermal roof of the yolk sac, beginning as a simple tubular shaped organ suspended by one common mesentery. The embryological development is completed through a complex set of rotations, after which it settles at the posterior abdominal wall.[1,2] Each part of the large intestine receives its blood supply according to its embryological development (midgut vs. hindgut), with the lymphatic drainage following the arterial blood supply (Figure1).

The intestinal wall is composed of four layers: (i) mucosa, (ii) submucosa, (iii) muscularis, and (iv) serosa. The innermost layer (the mucosa) consists of a single layer of cells (the colonic epithelium) which undergoes continuous renewal.[1] The glandular epithelium forms cylindrical crypts. The bottom of the crypts consists of proliferative and undifferentiated cells, which undergo differentiation into mature colonic epithelial cells at the top of the crypts. The fundamental pathway of colonic differentiation involves several complex intercellular signalling systems.[3] A detailed description of these systems is beyond the scope of this thesis. However, a large proportion of the genetic and epigenetic alterations in CRC affect embryological signalling pathways (e.g. *APC*, *BMP3*), displaying the interplay between embryological colonic development and CRC.

1.1.2. FUNCTION OF THE LARGE INTESTINE

Whereas almost all digestion and absorption of dietary nutrients occurs in the small intestine, the main function of the large intestine is the reabsorption of water and electrolytes. In this regard, the colon also absorbs some of the vitamins produced by intestinal bacteria (importantly menaquinones). It also stores the compacted faecal matter before evacuation during defaecation.

The large intestine also houses the largest amount of intestinal bacteria with more than 700 different species represented. The different microbes number more than 100 trillion weighing approximately 200 grams. The bacterial diversity differs to that of other parts of the intestine.[4] Most of these bacteria are commensals, meaning that they do not infer disease. However, in relation to other disease entities, or to antibiotic treatment, these bacteria may become virulent (e.g. *clostridium difficile*). Moreover, changes in the intestinal microbiome, have been linked to different disease entities (e.g. ulcerative colitis, Crohn's disease).[5–7] It is also believed, that these microbes and their metabolic output play an intricate part in CRC development, perhaps by influencing inflammation, DNA damage, and apoptosis.[8]

1.1.3. EPIDEMIOLOGY AND RISK FACTORS

Globally, CRC is the third most common cause of cancer (1.3 million/year) and the fourth leading cause of cancer related deaths (700,000/year).[9] Denmark, like most other western European countries, has a high incidence - roughly 70-90 cases per 100,000 persons per year.[10] The incidence in Denmark has risen in recent years, in part due to the introduction of the CRC screening programme in 2014 (described in detail below). In other high-income countries like the USA, the incidence has stabilised or started to decrease. This could be due to various screening programmes or colonoscopy with the removal of premalignant lesions, however, the decrease could also be attributed to changes in lifestyle and other unknown factors .[11]

CRC is a multi-factorial disease without one sole factor accounting for the majority of cases. Age is the most predominant risk factor, with more than 90% of CRC cases being diagnosed in patients above 50 years of age. In Denmark, the median age for CRC diagnosis is 72 years for colon cancer and 70 years for rectal cancer (Figure 2).

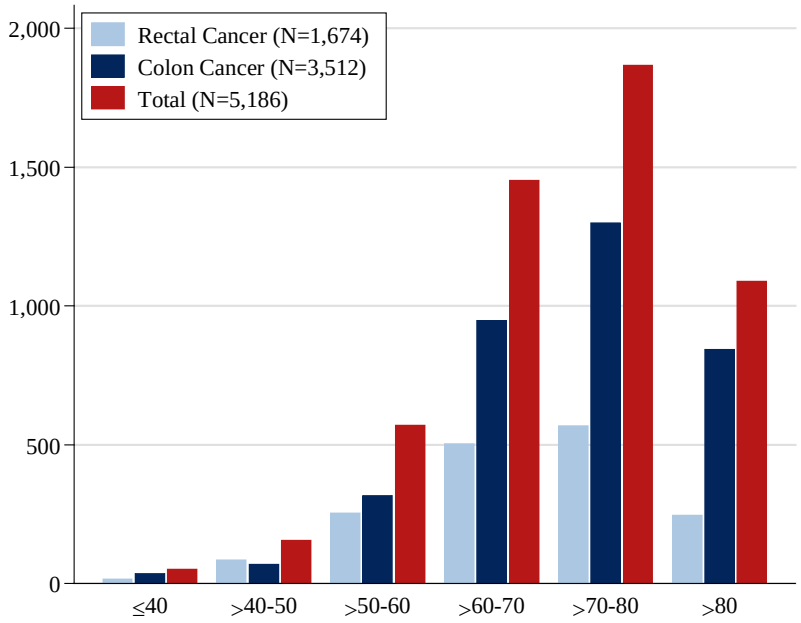


Figure 2 The annual incidence of colorectal cancer in Denmark (2014)
The number of colorectal cancer cases according to age in years

The established preventive factors include physical activity, acetylsalicylic acid, hormone replacement therapy, and colonoscopy with the removal of precancerous lesions. Physically active people have a 27% lower risk of proximal colon cancer (RR

= 0.73, 95%CI [0.66-0.81]) and a 26% lower risk of distal CRC (RR = 0.74, 95%CI [0.68, 0.80]), indicating no association by CRC subsite.[12] Ever usage of either oestrogen therapy (RR = 0.79, 95%CI [0.69, 0.91]) or oestrogen-progesterone therapy (RR = 0.74, 95%CI [0.68, 0.81]) has been associated with a decreased CRC risk, however the former use of oestrogen therapy infers no risk reduction (RR = 0.86, 95%CI [0.67, 1.11]).[13] Colonoscopy has been associated with a 77% lower risk of CRC (OR = 0.23, 95%CI [0.19-0.27]), with the largest reduction in risk being for left sided tumours (OR = 0.16, 95%CI [0.35, 0.20]) compared to right sided tumours (OR = 0.44, 95%CI [0.35, 0.55]).[14] Flexible sigmoidoscopy on its own also infers a benefit with a reduced risk of left sided tumours (RR = 0.67, 95%CI [0.59, 0.76]) and a decrease in CRC related mortality (RR = 0.72, 95%CI [0.65, 0.76]).[15] The effect of once-only sigmoidoscopy is emphasised in a incidence reduction of all CRC by 23% (HR = 0.77, 95%CI [0.70, 0.84]) and a reduction in CRC related mortality by 31% (0.69, 95%CI [0.45, 0.72]).[16]

Table 1 Risk and preventive factors for colorectal cancer

Sociodemographic factors	
Older age	↑↑↑
Male sex	↑↑
Medical factors	
Family history	↑↑
Inflammatory bowel disease	↑↑
Diabetes	↑
Helicobacter pylori infection	(↑)
Other infections	(↑)
Colonoscopy	↓↓
Hormone replacement therapy	↓
Acetylsalicylic acid	↓
Statin	(↓)
Lifestyle factors	
Smoking	↑
Excessive alcohol consumption	↑
Obesity	↑
Physical activity	↓
Dietary factors	
High consumption of red and processed meat	↑
Fruit and vegetables	(↓)
High consumption of fibre and whole grain	(↓)
Fish	(↓)
Dairy products	(↓)

Note: ↑↑↑ = very strong risk increase. ↑↑ = strong risk increase. ↑ = moderate risk increase. ↓↓ = strong risk reduction. ↓ = moderate risk reduction. Parentheses indicate probable but not established associations. Adapted from Brenner et al. (2014).[17]

Other lifestyle factors associated with decreased risk of CRC and CRC related deaths include: low alcohol consumption, low dietary intake of red or processed meat, low dietary fat intake, and high dietary fibre intake (Table 1).[18–23]

Approximately 15-35% of CRC patients have a family history of colorectal neoplasms, with known causative germline alterations being present in ~5%.[24] Genome wide association studies (GWAS), have identified a number of common single-nucleotide polymorphisms (SNPs) associated with increased CRC risk, however, the known SNPs only account for less than 7.42% of CRC heritability.[25] The search for other genetic risk factors accounting for the large heritability in CRC is therefore ongoing. The inherited syndromes of CRC with causative genetic alterations can broadly be classified into three groups based on clinical presentations: (i) Adenomatous polyposis syndromes, (ii) Nonpolyposis syndromes, and (iii) Hamartomatous polyp syndromes, each with distinct genetic alterations.[26] The inherited syndromes are more frequent in the younger population, and have varying disease penetrance. Hereditary nonpolyposis colorectal cancer (HNPCC) infers a lifetime risk of approximately 50-60% compared with ~100% in Familial adenomatous polyposis (FAP).[24] The molecular basis of the inherited syndromes of CRC is described in more detail below.

1.1.4. PROGNOSIS

The prognosis of CRC relates to the underlying cancer biology and in patient comorbidity. However, stage at the time of diagnosis remains the most important prognostic factor. The relative five-year survival for patients diagnosed with CRC in the U.S. from 2001 to 2007 was 90.1% in patients with localised disease, 69.2% in patients with locoregional spread, and 11.7% in patients with disease dissemination.[27] In Denmark, the survival is very similar, albeit somewhat lower, with an absolute five-year survival of 84.0% in early stage CRC patients and 13.0% in patients with distant metastases.[28]

The incidence of CRC and the CRC related mortality has dropped in recent years (Figure 3). This decrease can be attributed to four main categories. First, current treatments have improved compared to twenty years ago. Second, patients are increasingly aware of the early signs of CRC. At the same time, widespread availability and use of endoscopy has lowered the threshold for examination, with earlier diagnosis and polypectomy leading to both a decrease in incidence and mortality. Third, there could be fewer cases of CRC occurring in the first place, perhaps due to a reduction in the numerous risk factors listed above (Table 1). Finally, a large proportion of the reduced risk and mortality should be attributed to CRC screening programmes.

While the mortality and incidence have seen a steady decrease in the U.S. and other western countries, CRC incidence rates have increased in Denmark along with other countries.[29] CRC therefore remains a large disease burden, both on an individual level, but also with regard to the direct costs of medical care and lost productivity.[30]

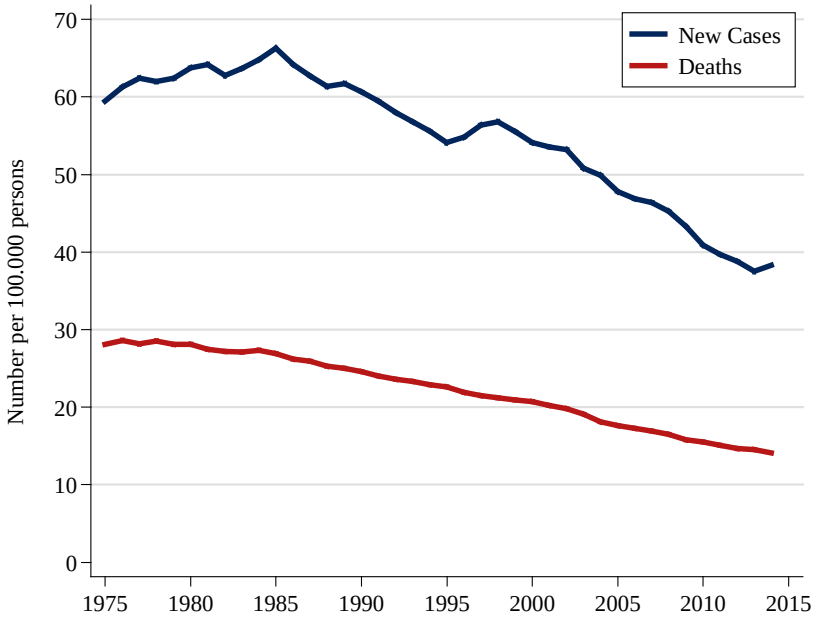


Figure 3 Number of unadjusted annual new cases and deaths of CRC in the U.S.
Data from the Surveillance, Epidemiology, and End Results (SEER) Program 9 (https://seer.cancer.gov/statfacts/html/colorect.html).[31]

1.1.5. SCREENING

Given that CRC develops over a relatively long time-period (~10 years), the high incidence rate, lack of early stage symptoms, increased risk by age, and potential cure with early stage detection, makes CRC ideal for population based screening. Currently, there are two main strategies for CRC screening: (i) sigmoidoscopy, and (ii) faecal occult blood testing. A Cochrane meta-analysis concluded, that both methods reduce the CRC specific mortality without a reduction in all-cause mortality, promoting the usage of both methods.[32] The advantage of stool-based screening is that it is cheap and non-invasive. While endoscopy-based screening is an invasive procedure, it offers the potential for intervention with the removal of pre-malignant lesions. In Denmark, a positive stool test is followed by colonoscopy.

In Denmark, CRC screening was introduced in 2014. The introduction of the screening programme was based on the studies above, and a feasibility study, which suggested improved survival in CRC patients diagnosed in the screening population when compared to non-responders.[33]

The national screening programme relies on the immunochemical faecal occult blood test (iFOBT), which has been preferred over the guiac-based test, due to increased sensitivity and specificity, improved patient compliance, and the lack of dietary restrictions.[34,35] The introduction of the Danish screening programme led to an instant increase in CRC cases by ~20% compared to the years preceding screening.[10,36] This is in sharp contrast to the decreasing incidence rates from the U.S. (Figure 3).

The adherence to CRC screening programmes are, however, not optimal. According to the National Health Interview Survey of 1987, only 23% of the U.S. population had recently been screened. In 2005 this proportion was only increased to 50%, displaying the need for improved screening methods.[37] This problem is further underlined by a Canadian study, suggesting that CRC screening adherence rates may be as low as 25% for stool based screening methods.[38] According to the annual report from the Danish Colorectal Cancer Screening Database adherence for CRC screening was 63% (95%CI [63, 64]) through the first 22 months after the initiation of the screening programme.[39] While the Danish compliance is superior to the adherence rates achieved in North America, the goal is still to increase compliance to at least 65%. In the U.S., the goal is even more ambitious. The goal is to reach an 80% adherence rate by 2018, especially by increasing access and awareness. This could possibly lead to a decrease in new cancer cases by ~280.000 and averting ~200.000 cancer deaths by 2030, showing the massive impact of compliance on screening programmes.[40]

The implementation of other less invasive methods could also increase CRC screening compliance.

1.1.6. PATHOLOGY

Neoplasms are an abnormal growth of tissue, usually not containing cysts or liquid areas. These tumours may be benign or malignant, depending on origin, the presence of cellular dysplasia, and growth patterns. In general, tumours derived from the epithelial cell-lining of the large intestine can be divided into two groups; adenomas and colorectal adeno-carcinomas. Other tumours of the large intestine account for less than 5%. The focus for the current thesis is the colorectal adenocarcinoma.

The pathological description of colorectal neoplasms relies on the morphology of the colonic epithelial cells. Immunohistochemistry is utilised in order to investigate some of the underlying molecular abnormalities in especially CRC. The structural abnormalities in colorectal neoplasia, will be described in the following, and the underlying molecular mechanisms is described in the section on molecular characteristics.

ADENOMA

The colorectal adenoma is the most frequent form of colorectal neoplasia in the western world. The frequency of adenomas and the incidence of CRC is directly proportional. In countries with a low risk of CRC, the incidence of colorectal adenomas is also low. Factors predisposing adenomas are the same, as the ones listed for CRC.

The development of adenomas involves dysregulation of proliferation patterns in the normal mucosa. The result is increased mitotic activity accompanied by decreased apoptotic capacity, resulting in an increased cell-number. The minimal lesion is the single crypt adenoma, which originates in so-called aberrant crypt foci.[41] These lesions are highly unstable and may develop into adenomas. Conventional adenomas are classified as tubular, tubulovillous, and villous, based on their structural architecture (Figure 4).[42]

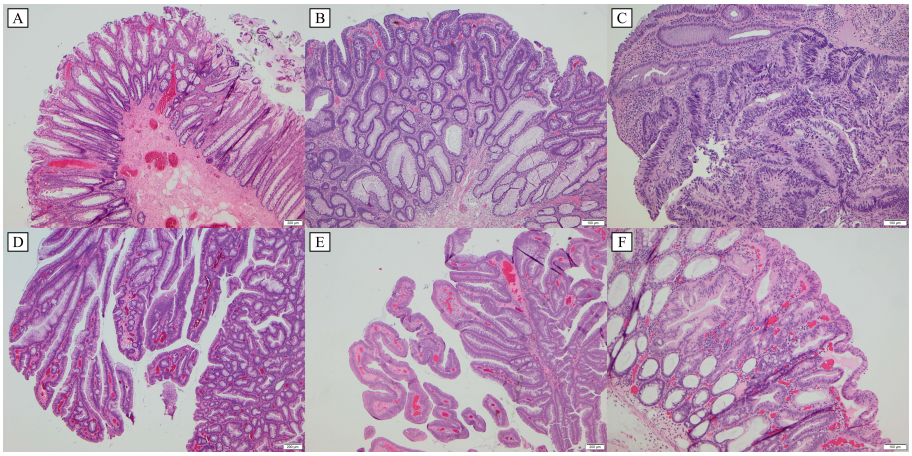


Figure 4 Classification of colorectal adenomas

A) Hyperplastic polyp. B) Low grade tubular adenoma (simple crypt-like dysplastic glands with < 25% villous component). C) High grade tubular adenoma (simple crypt-like dysplastic glands with < 25% villous component). D) Tubulovillous adenoma (intermediate lesions with 25-75% villous component). E) Villous adenoma (simple crypt-like dysplastic glands with > 75% villous component). F) Serrate adenoma (adenoma showing sawtooth or stellate (serrate) architecture). Christian Thomsen has kindly provided the images.

Every adenoma (with the exception of the hyperplastic polyp) has the potential for malignant transformation with an annual conversion rate of 0.25%.[43] An average adenoma patient is therefore only at moderate risk of CRC. Increase in size is closely related to increased cellular dysplasia making adenoma size an important risk factor for malignant transformation. The rate of conversion is highest in patients with adenomas of villous structure and/or severe dysplasia, with annual conversion rates of 17% and 37% respectively.[43]

ADENOCARCINOMA

More than 90% of CRCs are adenocarcinomas originating from the epithelial cells of the colorectal mucosa. Other types of CRC are metastases, neuroendocrine tumours, melanomas, with squamous cell, adeno-squamous cell, spindle cell, and undifferentiated carcinomas.[44] The sporadic adenocarcinoma is the focus of this thesis, and the other types will not be described further. Macroscopically, the CRC tumour can take one of four forms, with Type 4 being the least aggressive (Figure 5).

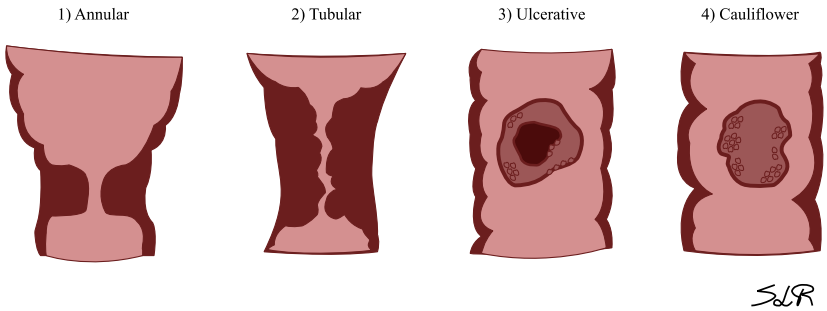


Figure 5 The common macroscopic variants of colorectal adenocarcinoma

Conventional adenocarcinomas are characterised by glandular formation and grade accordingly; well differentiated (>95% glandular formation), moderately differentiated (50-95% glandular formation), and poorly differentiated (<50% glandular formation).[44] Furthermore, they are also classified according to cellular morphology: (i) Mucinous adenocarcinomas; characterised by >50% extracellular mucin, (i) signet ring cell adenocarcinomas; characterised by >50% tumour cells with signet cell features (these are by definition poorly differentiated), and (iii) medullary adenocarcinomas; characterised by sheets of neoplastic cells with large nuclei and abundant cytoplasm (often poorly differentiated). Lastly, immunohistochemical staining is conducted for *CK20*, *CK7*, *CDX2*, and mismatch repair deficiency.

Most colorectal adenocarcinomas are moderately differentiated (70%) (Figure 6). Well and poorly differentiated adenocarcinomas account for approximately 10% and 20% respectively. Low grade tumours include the moderately and well differentiated adenocarcinomas. The determination of tumour grade is inherently a subjective exercise, making the reproducibility low, even among experienced pathologists.

The pathological report includes the histology of the resected adenocarcinoma, with a gross description, histologic type, tumour grade, tumour growth, lymph node involvement, blood vessel invasion, lymphatic invasion, perineural invasion, resection margins following surgery, and so forth.[44]

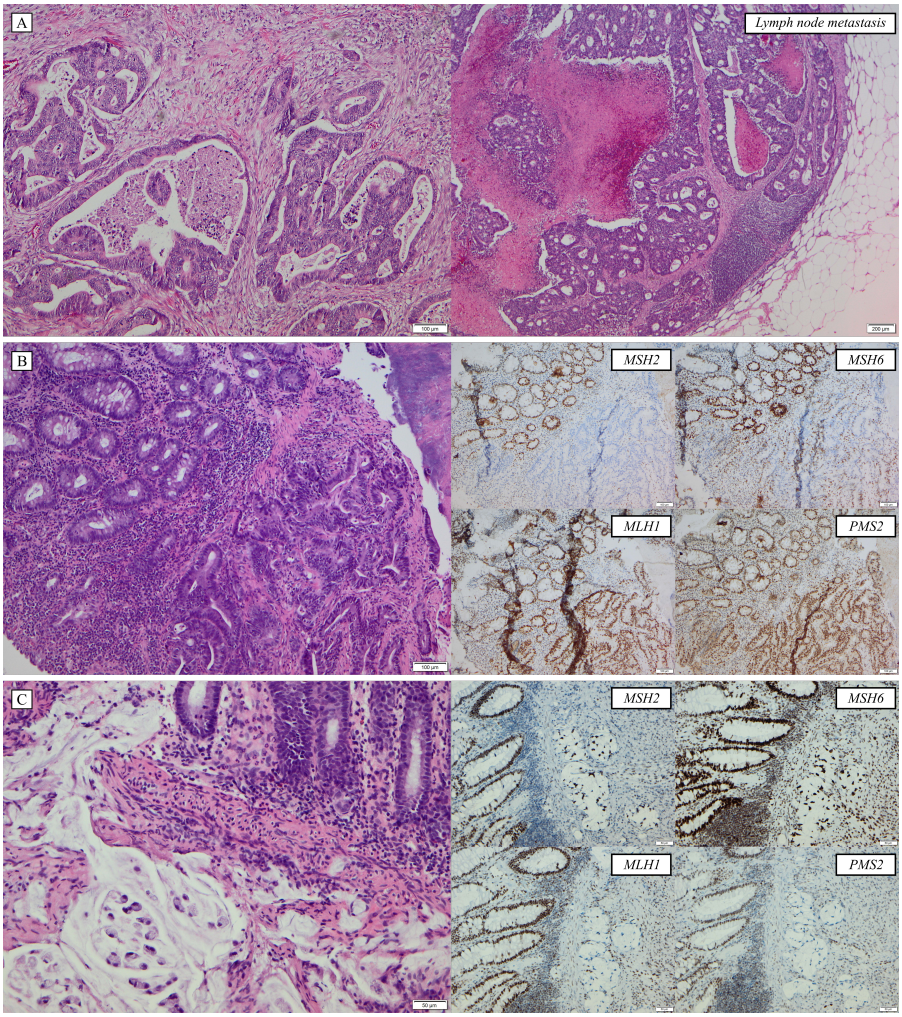


Figure 6 Classification of colorectal adenocarcinomas

A) Moderately differentiated adenocarcinoma with a lymph node metastasis. B) Mucinous adenocarcinoma with selective loss of *MSH2* and decreased expression of *MSH6*. C) Signet ring cell adenocarcinoma with decreased expression of *MLH1* and *PMS2*. Christian Thomsen has kindly provided the images.

1.1.7. DIAGNOSIS AND STAGING

The gold standard for CRC diagnosis remains complete colonoscopy with biopsy-verification.[45] In Denmark, patients are either referred for colonoscopy due to a positive screening test, or due to symptoms suggestive of lower GI malignancy. The main symptoms are those associated with increasing intestinal obstruction, including lower

abdominal pain (often colicky in nature) and abdominal distension. Alternating diarrhoea and constipation may be associated symptoms. Ulcerative tumours and tumours in the left colon may present with visible lower GI bleeding. Moreover, patients may have anaemia-associated symptoms in the case of visible of occult GI haemorrhage (e.g. pallor, shortness of breath, and fatigue). A somewhat large proportion debut with symptoms of metastatic disease (e.g. jaundice, ascites, and hepatomegaly). Weight loss is also a common unspecific symptom.

After CRC is diagnosed, subsequent clinical examination, laboratory tests, and imaging are used to detect or exclude metastatic disease. The current imaging modalities includes computed tomography (CT), positron emission tomography / computed tomography (PET/CT), and magnetic resonance imaging (MRI). CT of the thorax and abdomen is used as the primary tool for estimating the extent of tumour formation and importantly the detection of possible synchronous lesions, and/or metastases. However, MRI and CT are equally accurate, and MRI may be more beneficial in determining the involvement of the peritoneal surface or the involvement of regional organs - which is utilised in rectal cancer. PET/CT may be used in patients, where distant metastases are strongly suspected and not found by conventional CT imaging, but it is not recommended in the initial staging protocol.

The current staging system relies on the assessment of the anatomic extent of the disease at the time of diagnosis (Table 2). CRCs are staged according to the tumour-node-metastasis system (TNM).[46,47] The TNM classification is based on the clinical and radiological findings of tumour size and degree of direct invasion (T), involvement of locregional lymph nodes (N), and the presence of distant metastases (M). Refinement of the staging system, has led to the implementation of several, more complex tools. First, it includes a full stratification of the intestinal wall involvement and the peritoneal serosa, considering the number of regional lymph nodes involved. Second, it is a multidisciplinary tool, incorporating both clinical (pre-treatment classification: cTNM) and pathologic (postsurgical classification: pTNM, and after preoperative chemoradiotherapy: ypTNM) staging approaches. One unique feature of the CRC staging system is the in-situ cancer (Tis) which includes stromal invasion of malignant cells through the lamina propria and into, but not through, the muscularis mucosa. This is justified because the colorectal mucosa (unlike the mucosa elsewhere in the GI tract and other organs) lacks stromal lymphatics, making lymphatic metastasis impossible.

Tumour grade (or the grade of differentiation) has previously been recognised as a distinct prognostic factor. However, it has not been incorporated in the current CRC staging systems because of the lack of consensus regarding stratification, and the significant degree of interobserver variability.[48] Consequently, tumour grade was recently removed from the risk factors associated with stage II tumours (Figure 7).

Table 2 Colon and rectum cancer staging

Primary Tumour (T)				
Tis	Carcinoma in situ: Intraepithelial or invasion of the lamina propria			
T1	Tumour invades the submucosa			
T2	Tumour invades the muscularis propria			
T3	Tumour invades through the muscularis propria into the pericolorectal tissues			
T4a	Tumour penetrates to the surface of the visceral peritoneum			
T4b	Tumour directly invades or is adherent to other organs or structures			
Regional Lymph Nodes (N)				
N0	No regional lymph node metastasis			
N1a	Metastasis in one regional lymph node			
N1b	Metastasis in 2-3 regional lymph nodes			
N1c	Tumour deposit(s) in the subserosa, mesentery, or non-peritonealised pericolic or perirectal tissues			
N2a	Metastasis in 4-6 regional lymph nodes			
N2b	Metastasis in 7 or more regional lymph nodes			
Distant Metastasis (M)				
M0	No distant metastasis			
M1a	Metastasis confined to one organ or site (e.g. liver, lung, ovary, non-regional lymph node)			
M1b	Metastases in more the one organ/site or the peritoneum			
Anatomic Stage				
	T	**N**	**M**	**5-year survival rate**
Stage 0	Tis	N0	M0	-
Stage I	T1-2	N0	M0	86.3%
Stage IIA	T3	N0	M0	79.7%
Stage IIB	T4a	N0	M0	71.3%
Stage IIC	T4b	N0	M0	73.8%
Stage IIIA	T1-T2	N1/N1c	M0	85.4%
	T1	N2a	M0	
Stage IIIB	T3-T4a	N1/N1c	M0	70.3%
	T2-T3	N2a	M0	
	T1-T2	N2b	M0	
Stage IIIC	T4a	N2a	M0	46.9%
	T3-T4a	N2b	M0	
	T4b	N1-N2	M0	
Stage IVA	T1-4	Any N	M1a	15.2%
Stage IVB	Any T	Any N	M1b	9.7%

Note: The staging system is according to the American Joint Committee on Cancer (AJCC) 7th Edition staging system.[46] The 5-year survival estimates according to stage are retrieved from Lan et al. (2012).[49]

1.1.8. TREATMENT

The only curative treatment for CRC is total resection of the affected colorectal segment along with all associated locoregional lymph nodes, which could potentially harbor lymph node metastasis. Current treatment guidelines involve neo-adjuvant chemo- and radiotherapy, surgical resection, and adjuvant chemotherapy. All patients are discussed in a multidisciplinary conference comprising surgeons, pathologists, radiologists, and medical oncologists. The basis for the optimal patient treatment relies on pre-treatment staging and the distinction between colon and rectal cancer. The standard treatment of colon cancer is surgery, combined with adjuvant chemotherapy depending on the stage of disease at the time of diagnosis (Figure 7). The treatment options for rectal cancers and the treatment options for colon cancer are highly similar, with the addition of neoadjuvant chemoradiotherapy. The difference in treatment between rectum and colon cancer also relies on the difference in embryological origin, with the vascular drainage from the rectum going straight to the vena cava, and the vascular drainage from the colon going to the portal circulation, rendering pulmonary metastasis more common in rectal cancer.

COLORECTAL SURGERY

Principally, CRC surgery involves the resection of the tumour-bearing bowel segment, with adequate proximal, distal, and lateral resection margins, central ligation and division of the supplying blood vessels, and removal of the attached mesocolon/mesorectum. This type of resection is based on the sharp dissection of distinct embryological planes to avoid breaching the visceral layer and the possibility for tumour spread to the peritoneal cavity.[50,51] The complete resection (R0) involves the aforementioned elements and complete *en-bloc* removal of the resected bowel.[52] Bowel-continuity is re-established by a sutured or stapled anastomosis. The presence of residual tumour is reported as either macroscopic (R2) or microscopic (R1).

If the circumferential resection margin involvement is below 1 mm, the patient has not received radical surgical treatment. If the involved lymph nodes (< 12 lymph nodes) are not removed, the resection is also considered incomplete. In both cases, post-operative staging cannot be considered optimal.

Laparoscopic surgery is the approach of choice in the case of uncomplicated cancer of the right and left colon, because it offers faster recovery and less morbidity compared to the open approach.[53–58] In the case of bulky and advanced colonic lesions, laparotomy and subsequent resection is still considered the treatment of choice.

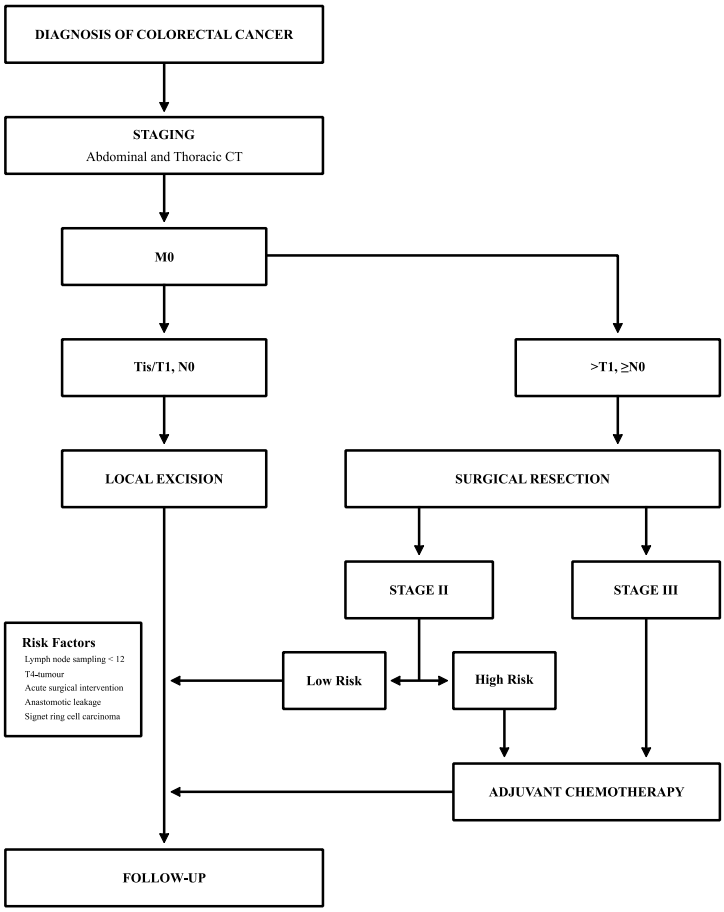


Figure 7 Treatment algorithm for non-metastatic colon cancer

The algorithm is in accordance with the National guidelines in Denmark provided by the Danish Colorectal Cancer Group (2017). The distinction between high and low risk stage II colon cancer is based on the risk of recurrence, and the associated risk factors are listed in the box "Risk Factors". Patients > 75 years, patients with WHO performance status > 2 due to comorbidity, or patients with MSI tumours should never receive adjuvant chemotherapy regardless of the presence of risk factors. Adjuvant chemotherapy is usually based on 5-Flourouracil, oxaliplatin, or capacitabine.

COLORECTAL ONCOLOGY

Medical treatment of CRC involves the use of cytostatic drugs to induce tumour regression. Current treatment regimens can only be used as supplements for curative surgery; otherwise, the usage is only palliative in nature.

Adjuvant chemotherapy for CRC without metastasis can commence within five weeks after surgery or when the patient is fit to receive therapy. The standard treatment for stage II CRC is Capecitabine or 5-fluorouracil in combination with folinic

acid (de Gramont). For stage III colon cancer, treatment is based on the combination of folinic acid, 5-fluorouracil, and oxaliplatin (FOLFOX), or a combination of capecitabine, and oxaliplatin (CAPOX).[59,60] For stage III rectal cancer, the choice of treatment depends on the rectal tumour location, with one of the treatment options being neo-adjuvant chemo-radiotherapy.[61]

For metastatic CRC, treatment is tailored according to the degree of disease dissemination, a pivotal factor being the resectability of the metastasis.[62] Resection or radiofrequency ablation (RFA) of solitary liver metastasis provides a strong survival benefit in these select patients.[63,64] Moreover, molecular biomarkers are utilised in the choice of treatment with monoclonal antibodies. One example being tumours harbouring an activating mutation in *KRAS* or *NRAS*, in which patients will not derive benefit from treatment with epidermal growth factor (EGFR) monoclonal antibodies such as cetuximab and panitumumab.[62] Other types of molecular changes utilised in the treatment of metastatic CRC is described in the section on molecular characteristics, below.

FOLLOW-UP

The aim of follow-up is the detection of local or distant recurrence, metachronous colonic lesions, and assessment of late post-treatment morbidity. Intense follow-up programmes have led to an improved overall survival by 7-13%.[65] However, there is still some ambiguity regarding the optimal structure of CRC follow-up programmes, and the results of the COLOFOL trial is therefore much anticipated.[66] The Danish follow-up programme includes colonoscopy three months after surgery (if colonoscopy was not conducted prior to resection), followed by colonoscopy every five years postoperatively (until the age of 75). CT of the thorax and abdomen should be conducted one and three years after diagnosis.

Preoperative carcinoembryonic antigen (CEA) levels are analysed and utilised as the only blood-based biomarker for the detection of completeness of surgery (increased preoperative CEA-levels, should normalise after surgery). An increase in post-operative CEA-levels is used as an indication of disease recurrence, however, with lacking sensitivity.[67] The limited sensitivity of CEA makes it unfit for both CRC detection and screening.

1.2. MOLECULAR CHARACTERISTICS

The molecular basis for cancer in general, is the inactivation of genes associated with cellular regulation and the activation of genes involved in increased cell survival and metastatic properties.

Theodor Boveri already proposed the concept of cancer being a genetic disease in somatic cells in 1914, prompted by the works by David von Hansemann.[68] Their pioneering efforts were based on aberrant mitosis, chromosome numbers and abnormalities in centrosomes, and Boveri even anticipated the era of tumour suppressor genes and oncogenes. In 1971, Alfred G. Knudson formulated his *two-hit hypothesis*, considering cancer as a multi-mutational disease, with inherited forms already bearing one of the two hits necessary for cancer formation further emphasising cancer as a genetic disease.[69] The discovery of proto-oncogenes in cancer cells by Dominic Stehelin, J. Michael Bishop, and Harold E. Varmus provided the evidence, that alterations in our own genes can cause cancer formation – the basis for J. Michael Bishop and Harold E. Varmus being awarded the Nobel Prize in 1989.[70]

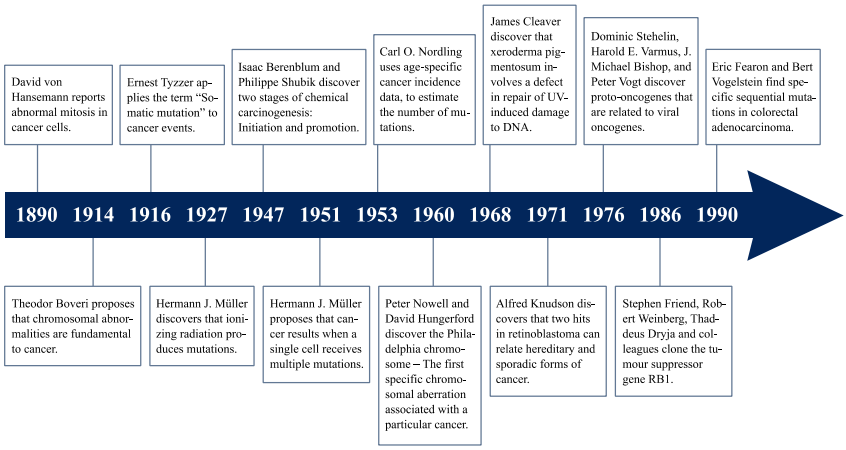


Figure 8 Timeline of the discovery of cancer related genetic alterations
A broad overview of the most important findings associating mutational events with carcinogenesis. Adapted from Knudson (2001).[71]

Today, we know that tumours develop from benign to malignant lesions by acquiring a series of mutations over time. This multi-step carcinogenesis paradigm relates well to the fact that cancer takes several years, even decades, to develop making age, the most pertinent risk factor.

Genome-wide data have shown, that a typical colorectal tumour harbours approximately 55-110 mutations.[72] The first genomic aberration is called the *gate-keeping*

mutation, which provides a selective growth advantage for the tumour cell compared to the surrounding normal epithelial cells. The mutations, which confer a selective growth advantage, are called *drivers*, and to date, approximately 120-140 have been discovered (~65 oncogenes and ~75 tumour suppressors genes).[72,73] Oncogenes are usually affected by missense mutations, altering amino acid residues, leading to a gain-of-function or non-regulated protein expression (e.g. *KRAS*, *BRAF*, and *SMAD4*). Tumour suppressor genes are typically affected by any kind of mutation in the coding region, which ultimately disables the protein product (e.g. *RB1*, *TP53*, and *APC*). The *drivers* can be classified into one or more of three distinct cellular processes: (i) cell fate, (ii) cell survival, and (iii) genome maintenance.

Alterations in genes involved in the above mentioned cellular processes, will always produce a cancer cell with a selective growth advantage over normal cells. Several different genes can result in the same growth advantage, with a large overlap between the different cellular processes, making the identification of *drivers* difficult. Although many *drivers* have been identified, the vast majority of mutagenic alterations in human cancers are bystanders. These *passengers* do not contribute to cancer formation, instead, they constitute the bulk of the somatic alterations in each individual tumour. The strict distinction between molecular *drivers* and *passengers* remains enigmatic, and the subject for vigorous research.

1.2.1. THE GENETIC BASIS FOR COLORECTAL CANCER

The molecular characteristics of CRC initiation and progression through the adenoma-carcinoma sequence, was first described by Eric R. Fearon and Bert Vogelstein in 1990.[74,75] The adenoma-carcinoma sequence describes the stepwise progression from normal to dysplastic epithelium, along with the accumulation of several genetic alterations (Figure 9). This accumulation of gene mutations is non-random and initiates colorectal carcinogenesis through the deregulation of pathways that modulate cellular differentiation, proliferation, and apoptosis.[72]

All CRCs harbour genetic alterations, including single base substitutions, and larger structural variations (e.g. aneuploidy).[76,77] A detailed description of all the mutational events related to CRC is beyond the scope of this thesis; however, some of the most important genetic aberrations will be described below.

The classical description of the multistep genetic model for CRC starts with the inactivation of *APC*. *APC* mutation occur in more than 80% of sporadic CRCs.[78] *APC* germline mutation is also responsible for the autosomal dominant inherited FAP syndrome associated with multiple colorectal adenomas.[79] *APC* inactivation is followed by two other frequent mutations, *KRAS* and *BRAF*. Mutation of *KRAS*/*BRAF* leads to an enhancement of the gene product, stimulating cellular growth through the

MAP/ERK pathway. *KRAS* is mutated in 33% and *BRAF* in 10% of CRCs.[80,81] The *KRAS*/*BRAF* protein is downstream of EGFR, making treatment with monoclonal antibodies against EGFR redundant in mutant tumours. Other additional mutations resides in the TGF-β (*SMAD4*), PIK3CA, and TP53 pathways along with loss of heterozygosity and aneuploidy.[82] Indeed, these associations of chromosomal instability are seen in ~85% of invasive CRC.[77] Other molecular alterations are simplistically illustrated in Figure 9.

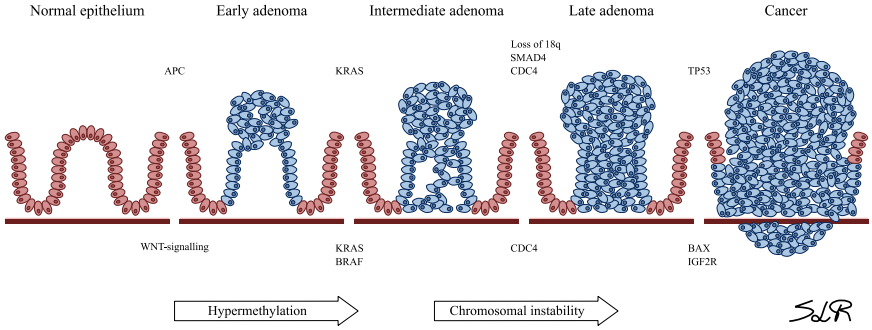


Figure 9 The adenoma-carcinoma sequence

Colorectal carcinogenesis develops through a series of genetic alterations, involving oncogenes (e.g. *BRAF*, *CDC4*, *KRAS*, and *SMAD4*) and tumour suppressor genes (e.g. *APC*, *BAX*, and *TP53*). The chromosomes most frequently deleted include 5q, 17q, and 18q.

CRCs harbouring a deficient mismatch repair (MMR) system regardless of origin account for 15-20% of CRC cases.[83] The MMR system is responsible for the surveillance, and correction of errors introduced in microsatellites (interspersed repetitive elements throughout the human genome – prone to mutations). The main MMR genes are *MLH1*, *MSH2*, *MSH6*, and *PMS2*, and downregulation of either of these, will infer microsatellite instability (MSI). HNPCC patients with Lynch syndrome have an autosomal dominant germline defect in one of the MMR genes causing their predisposition to CRC (described above).[84] CRCs with MSI are associated with right-sided tumours, with an improved prognosis compared to sporadic tumours.[85,86] Moreover, MSI CRCs are more susceptible to treatment with PD1-antibodies.[87] Modes of MSI also involve epigenetic inactivation of *MLH1* via promoter hypermethylation. This mode of tumour development due to hypermethylation is the focus of this thesis and described in further detail below.

CRCs arising in the proximal and distal colon appear to arise from distinct molecular pathways, reflecting the physiological differences between the two bowel segments. These molecular differences may explain differences in morphology (increased proportion of flat polyps in the proximal colon).

1.2.2. HYPERMETHYLATION AND COLORECTAL CANCER

Epigenetics is the study of heritable changes in DNA expression, without an underlying change in the DNA sequence. DNA methylation and histone modifications are the fundamental elements in epigenetics, and each function to adapt the chromatin, and hence genomic expression.[88] These changes are inherently reversible, and can be altered by multiple environmental factors and different disease states, making them an interesting field for the development of epigenetic therapies.[89] DNA methylation and histone modifications are closely linked molecular phenomena, however, only DNA promoter hypermethylation, will be described in the section below. Stephen B. Baylin and Peter A. Jones (2016), elegantly and thoroughly describe histone modifications and other epigenetic processes in relation to cancer.[89]

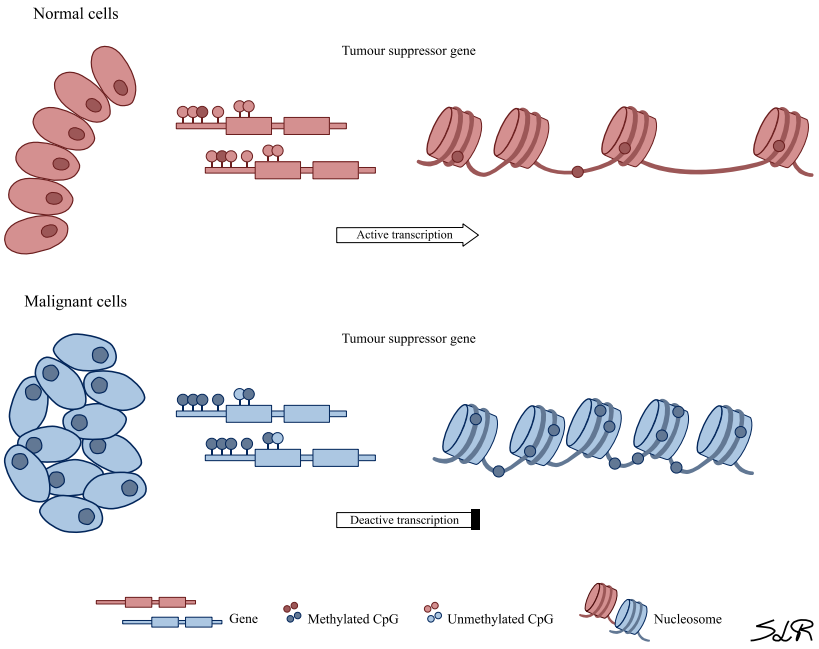


Figure 10 The effect of DNA hypermethylation on transcriptional activity

The promoter region of tumour suppressor genes is normally hypomethylated, and different histone modifications ensure a stable open chromatin structure (euchromatin) making transcription possible. In the malignant transformation of epithelial cells, the promoter regions become densely methylated, and different histone modifications makes the chromatine more tightly packed (heterochromatin) leading to transcriptional repression.

DNA methylation involves the addition of methyl groups to cytosine preceding a guanine (CpG). Cytosine methylation in the CpG dinucleotide motif is essential for normal mammalian development, X-chromosome inactivation, and genomic imprinting.[90] During embryological development, the DNA of the embryo is first subject to

an initial wave of demethylation, after which (the eight-cell stage) the embryo DNA receives a cascade of *de novo* methylation. The target for methylation is the 5'-carbon, and the following dinucleotide is denoted 5-methylcytosines. Methylation of the 5'-carbon is catalysed by one or more of the group of enzymes called DNA methyltranferases (DNMTs) using a methyl-group from S-adenyl methionine.[91] CpGs are found genome-wide, however, not at a frequency, which would otherwise be expected by chance. They often cluster in regions called CpG islands, and global CpG island hypomethylation and regional hypermethylation is part of nearly all human malignancies.[92,93] CpG island methylation at key sites in the DNA promoter regions have been shown to lead to decreased transcriptional expression, and thereby effectively hindering tumour suppressor function. This mechanism is thought to originate based on the blockage of transcription factors and conformational changes in the chromatin structure (Figure 10).[94,95] Several studies have suggested, this hypermethylation associated gene silencing, however, the association is not constant. Recent epigenome-wide association studies (EWAS) have shown, that the inverse correlation between DNA hypermethylation and gene expression might be as low as 40%.[96] Nevertheless, it is still widely accepted that promoter hypermethylation infers decreased transcriptional activity. Previous studies suggest that the regulation of DNA methyltransferase 1, which possesses *de novo* methylation activity for CpG-islands, becomes dysregulated during colorectal carcinogenesis.[97,98] However, increased promoter hypermethylation in cancer is not merely a result of a transcriptional deregulation in the three known DNMTs.[99] The underlying cause for deregulation in promoter methylation remains unknown.

The evaluation of gene hypermethylation status has revolved around tissue studies. Historically, the identification of putative methylation biomarkers has been conducted using candidate gene approaches. However, advancing genome-wide technologies have significantly increased the current knowledge on the scope of DNA methylation changes in CRC. In normal colonic cells, various patterns of DNA methylation exist based on anatomical location, embryological origin (midgut vs. hindgut), gender, and patient age.[100] Generally, DNA of these normal cells exhibit a low degree of methylation. Colorectal tumour cells are very different in this regard. These cells are characterised by methylation of various tumour suppressor genes and intragenomic hypomethylation of repetitive elements (e.g. *LINE1*).[101–103] Hypermethylation is an early event already seen in adenomas and aberrant crypt foci (e.g. *CDKN2A*) and associated with poor survival in stage II and III CRC.[104,105]

The number of methylated gene promoter regions related to CRC initiation and progression is enormous.[106] Tissue studies have suggested that CpG island hypermethylation might be associated with a distinct CRC subtype termed the *CpG island methylator phenotype* (CIMP).[107–109] CIMP has been characterised by a high frequency of methylation at key genomic sites, and associated with right sided tumours, a high degree of microsatellite instability, *BRAF* mutation, and a less aggressive pheno-

type.[110] CRC tumours with CIMP and tumours with a high degree of mutations (chromosomal instability), have been shown to be inversely correlated, leading to the understanding that there may be three distinct subtypes of CRC following three distinct molecular pathways; namely the chromosomal instability pathway, the MSI pathway, and the CIMP pathway. The definition of CIMP currently relies on the works by Weisenberger et al. (2006) who identified a robust five gene panel, currently used to define CIMP tumours.[108,111] However, the cause or the mechanism for CIMP has still to be identified, and one study has even postulated, that CIMP merely is a statistical artefact.[112] Moreover, a recent study found that CIMP did not have any impact on the survival of CRC patients.[113] The sensitivity of the panel therefore remains to be validated.[114]

The candidate gene-approach studies and EWAS show, that there is a marked difference in DNA promoter methylation between cancer types.[99] Each cancer type may therefore harbour its own hypermethylation signature. A recent study even showed, that CRCs harbour more aberrant DNA hypermethylation compared with somatic mutations.[115] Since methylation patterns are already altered in colorectal adenomas, this promotes their utilisation as biomarkers for CRC detection.[115,116] However, it is still not known, if DNA hypermethylation is a *driver* or *passenger* alteration in the initiation and progression of CRC.

1.2.3. THE CONSENSUS MOLECULAR SUBTYPES

The amount of data from recent GWAS and studies on genome expression, have increased in recent years. These studies have focused on improving our understanding of CRC, and primarily aimed at identifying molecular subgroups to enable improved diagnostics, stratification according to prognosis, and response to adjuvant chemotherapy. The combined effort to understand the molecular biology of CRC has led to the characterisation of CRC by means of the consensus molecular subtypes (CMS 1-4).[117] Each subtype has a set of distinct genomic and epigenomic traits. Moreover, their mode of pathway enrichment is relatively unique. Proximal tumours with CMS1 is characterised by hypermutation and hypermethylation, whereas CMS2-4 are more closely linked to an increase in chromosomal instability. These subtypes could potentially serve as adjuncts in treatment decision making. Proximal CMS1 and CMS3 tumours have a decreased expression of EGFR ligands, whereas CMS2 reveals an overexpression of EGFR ligand, prompting sensitivity for EGFR monoclonal antibodies (e.g. cetuximab). Even though it was developed using multiple datasets comprising data from 4.151 CRC patients, the consensus molecular subtypes, have yet to prove that they correctly stratify patients in clinically relevant subgroups. A recent study has even proposed another model (consisting of three cancer cell archetypes and five tumour archetypes) refining the characterisation of CRC.[118] This includes

the dARE archetype (a stratification of CMS2) which is associated with a higher bacterial read count than other chromosomally instable tumour samples. This illustrates, that characterisation of human tumours is complex, probably due to tumour heterogeneity, laboratory techniques, imperfect assays, and interpretation of the vast amounts of data. Whether or not these molecular characterisation patterns have any clinical utility, remains to be elucidated.

1.2.4. CIRCULATING MOLECULAR BIOMARKERS

Extracellular nucleic acids was first described by Mandel and Metais in 1948.[119] Subsequently, Leon et al. (1970) showed that the level of circulating cell-free DNA was increased in cancer patients.[120] Since then, elevated levels of circulating cell-free DNA have been observed in nearly all cancer types (e.g. breast, lung, and ovarian cancer), and a number of other non-malignant disease entities.[121,122]

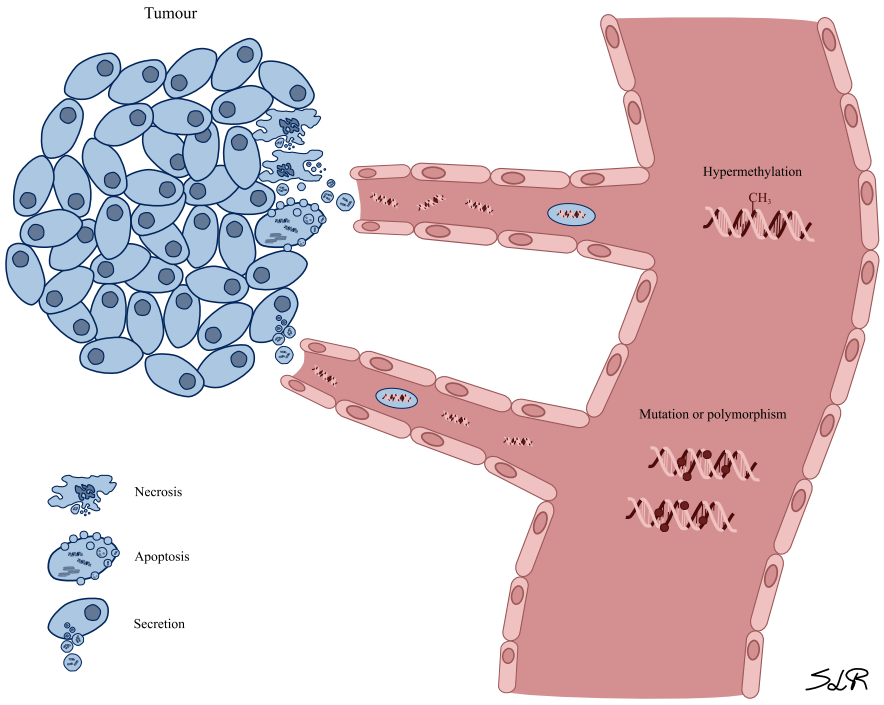


Figure 11 Circulating cell-free DNA

DNA is released by manner of necrosis, apoptosis, phagocytosis, and active secretion. DNA is shed as both single and double stranded DNA with an approximate length of 160-180 base pair. Circulating tumour DNA is even more fragmented with an approximate length < 145 base pair. Membrane *blebbing* leads to the release of circulating DNA in apoptosis. Circulating DNA related to exosomes are structurally associated with the membrane (> 2,500 base pair), or resides inside the vesicle (100-2,500 base pair).

Circulating cell-free DNA can be analysed using the same molecular methods, as DNA obtained through tissue biopsies. This makes it possible to detect various tumour related mutations, polymorphisms, and aberrant hypermethylations. One of the challenges in analysing cell-free tumour derived genomic alterations, is the scarcity of these fragments. The number of cell-free tumour DNA fragments has been shown to be less than 10 per 5 ml plasma for stage I CRC, and detecting these fragments in a larger amount of normal DNA is inherently difficult.[123]

DNA from malignant cells is highly fragmented with a molecular size smaller than 145 base pairs. Assays used for the detection of these small amounts of DNA needs to be highly sensitive in order to detect the mutational changes or various degrees of hypermethylation. Current methods for the DNA methylation analysis are described in the following section.

1.2.5. METHODS FOR HYPERMETHYLATION ANALYSIS

A large number of strategies for DNA methylation analysis exists, and have been used to describe different methylation biomarkers.[124] These methods can broadly be classified into four main categories:

i. Restriction enzyme digestion
ii. Affinity-based analysis
iii. Bisulphite modification
iv. Direct sequencing

The method of choice is very much dependent upon the hypothesis of the individual study being performed (e.g. the evaluation of putative disease methylation sites or discovery of novel areas of aberrant DNA methylation). This is due to the different strengths and weaknesses of each methodological. Table 3 provides a short description of the methods currently available for methylation analysis.

The cornerstone in the analysis and validation of aberrantly methylated DNA has been bisulphite conversion and subsequent analysis by polymerase chain reaction (PCR). This method relies on the differential reaction of methylated and un-methylated cytosine with the reagent. Un-methylated cytosine undergo deamination converting them into uracil, whereas, methylated cytosine residues remain intact, creating methylation-dependent sequence differences in the genomic DNA (Figure 12). Sequence variants at a particular locus can subsequently be analysed using PCR amplification targeting bisulphite converted DNA or microarray based technologies.

Table 3 Methods for DNA methylation analysis

	Description	*Advantages*	*Limitations*
Restriction enzyme digestion	Use methylation-sensitive enzymes, with different affinities for cytosine methylation status. The restriction enzymes, MspI and HpaII, both recognise the same sequence (CCGG), however, with their cleavage capabilities are modified by various cytosine modifications.	Methylation profiles are reproducible A relatively simple method	Some methods require radioactive material Labour intensive Limited to enzyme digestion sites Not well suited to distinguish moderately and weakly methylated fragments
Affinity based analysis	Use shearing to generate random DNA fragments followed by denaturation and immunoprecipitation of single stranded DNA fragments with monoclonal anti-5-methylcytosine antibodies. The isolated DNA fragments are then enriched and amplified using PCR, microarray, or sequencing based technologies.	Fast and reliable assessment of mean methylation levels of large DNA regions Can distinguish 5-methylcytosine from 5-hydroxy methyl-cytosine Reagents are commercially available and easy to use	Requires single stranded DNA Limited by the sensitivity and specificity of the antibody Provides no information on single nucleotide methylation Sequencing bias
Bisulphite modification	Use treatment with bisulphite to generate a difference in the DNA sequence. It relies on the differential reaction of methylated cytosines versus un-methylated cytosines with bisulphite. The conversion can be detected using various methods combined with PCR, microarray or sequencing based technologies.	Amenable for quantitative analysis Allows for global DNA methylation analysis Most widely accepted and utilised method When combined with PCR it is rapid, simple, and cheap	Significant degradation of DNA during conversion Incomplete conversion will lead to erroneous interpretations Cannot distinguish 5-methylcytosine from 5-hydroxy methylcytosine Requires sequencing based technologies for global DNA analysis
Direct sequencing	Use the time difference between base incorporations (the inter-pulse durations) or electrolytic current signals to detect aberrantly modified nucleotides.	Allows for global DNA methylation analysis Does not rely on bisulphite conversion	The accuracy for 5-methylcytosine may be limited, due to weak kinetic signals or low signal resolution Labour intensive Sequencing bias

Note: Description of the four overall methods for DNA methylation analysis. Most of the methods listed above are thoroughly reviewed by Olkhov et al (2012).[124] The direct sequencing method is described elsewhere.[125,126]

Methylation specific PCR is the most utilised method for locus-specific bisulphite modification analysis, with a reliable application in numerous clinical studies.[127–132] A number of variations for methylation specific PCR exist, each being suitable for a number of different research applications. In methylation specific PCR, the bisulphite modified DNA is subsequently amplified using primer sets and a fluorescent probe specific for the methylated sequence of interest. A range of modifications have been made to the methylation specific PCR technology, however, the overall principle remains the same.[133–138]

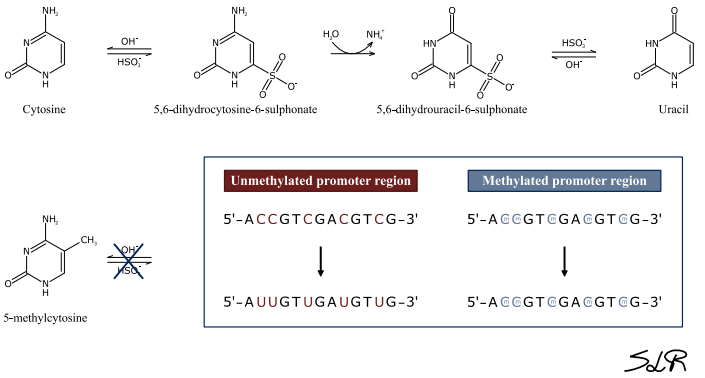


Figure 12 Bisulphite treatment of cytosine and 5-methylcytosine

Cytosine is converted into uracil through aromatic sulphonation and deamination. The 5-methylcytosines are not converted during bisulphite treatment. The bisulphite treatment creates a difference in the DNA sequence (box) amenable for detection by PCR based methods.

One of the major benefits of bisulphite based assays, is that the method only relies on small amounts of sample DNA, and the high-throughput capacity of most of the PCR based assays.[138]

One of the major drawbacks in bisulphite-based methods is the massive degradation of DNA rendering less DNA for subsequent PCR analysis. Moreover, incomplete conversion could lead to erroneous interpretations.[139] Another drawback in using PCR based methods is the loss of information, since only one CpG-island can be analysed at any one time. It is therefore not possible to make inference about the methylation status of the whole promoter region. Lastly, as with other techniques relying on PCR, false positives can arise if primers are of poor quality or used at incorrect temperatures.

The gold standard for evaluating DNA methylation remains bisulphite conversion followed by sequencing, because it enables the base-pair resolution of CpG methylation.[140] However, the procedure is costly and time-consuming, generating vast

amounts of data, necessitating complex bioinformatics (base-calling), making interpretation difficult.

The method utilised through this thesis, is based upon an optimised methylation specific PCR protocol described by our group.[141] The method relies on a rapid bisulphite conversion, rendering the method more readily applicable. Methods used in previous studies, and most commercially available kits are very time consuming, relying on treatment with bisulphite for 4-18 hours before subsequent PCR or sequencing analysis.[142] The method employed by our group takes advantage of deamination at a higher temperature to ensure complete bisulphite conversion with a yield of nearly 60% DNA post treatment.[143,144] A recent review of four commercial kits also suggests that our method is more than comparable as these methods produced a DNA recovery ranging from 33.2% to 55.0%.[145]

The method uses *MEST* as a reference gene for methylation. *MEST* is useful as a naturally occurring model system, due to the fact that *MEST* is maternally imprinted rendering only one allele methylated.[146] This type of hemimethylation makes it possible to ensure the correct assessment of the effect of methylation changes in other putative methylation biomarkers.[141] All primers and probes are specific for methylation and synthesised *in-house*. We used hypermethylation specific molecular beacons for the real-time PCR assay.

CHAPTER 2. AIMS AND SCOPE

This thesis is based on the epigenetic foundation from which cancer develops. The number of studies addressing molecular changes in human malignancies have exploded in recent years, and the characterisation of solid cancers in general and CRC in particular expands using improved molecular methods. Hence, the development of novel biomarkers for all types of human malignancies is an extremely rapid process. However, not many of these putative biomarkers are currently used in clinical practice. This lack of biomarker utilisation is primarily due to a lack of performance in subsequent validation studies or the scarcity of replication studies in general. Alas, the efficacy of cell-free DNA hypermethylation as biomarkers for CRC lacks sufficient evaluation.

The aims of the current thesis were:

1) To conduct a systematic review of cell-free hypermethylated DNA as a biomarker for CRC. The review would serve two purposes:

    i. To give a broad overview over the studies conducted so far.
    ii. To select hypermethylated regions for further analysis.

2) To evaluate the potential of hypermethylated cell-free DNA in plasma for CRC detection.

3) To evaluate the value of hypermethylated cell-free DNA in plasma for CRC prognostication.

CHAPTER 3. METHODS

The methods for each of the three aims listed above, will be described in detail in this section. First, a description of the selection of gene promoter regions to be evaluated as potential predictive biomarkers for CRC. Secondly, a description of the study populations for each study, followed by a description of the stepwise hypermethylation analysis. Lastly, the data handling and statistical analyses will be elaborated.

3.1. LITERATURE STUDY

The focus of this thesis was the evaluation of hypermethylation biomarkers for CRC detection and prognostication. These were selected through a comprehensive review of the current literature.

Eligible studies were all articles where cell-free DNA was analysed for methylation of putative CRC promoter regions. Cell-free DNA has been evaluated in both stool and blood samples, for the detection of CRC. We chose to include these stool based hypermethylation studies to make the most comprehensive analysis of hypermethylation biomarkers for CRC detection. The literature search also included all the studies on tumour biopsies to ensure the completeness of the literature search.

Embase, Web of Science, and Medline databases were searched, using the following terms: DNA methylation, CpG islands, tumour sample, stool sample, blood, serum, plasma, PCR, microarray analysis, genome sequencing, biopsy, and blood analysis (updated September 2015).

Studies were considered if they analysed one or more hypermethylated gene promoter regions from stool or blood samples. All studies were included, regardless of size or analytical method. Studies considering solely, tumour biopsies, cell lines, and animal models, were subsequently excluded along with studies written in languages other than English.

For the subsequent hypermethylation analysis we chose biomarkers with the capability to distinguish early stage CRC patients from healthy control individuals. Moreover, we included hypermethylation markers found in other types of solid cancers (e.g. pancreatic cancer). The names and known functions of the 30 genes analysed for promoter hypermethylation in cell-free plasma are listed in Supplementary table 1.

3.2. STUDY POPULATION

The thesis is based on the analysis of blood samples from a previous study conducted at the Department of Gastrointestinal Surgery, Aalborg University Hospital.

Between October 2003 and November 2005, 314 consecutive CRC patients were evaluated as participants in a study of pre- and postoperative venous thromboembolism. These patients were admitted for intended curative surgery and had blood samples drawn before any kind of treatment. Patients were excluded if they had any previous or concomitant cancer of any origin (within three years), known congenital thrombophilia, thromboembolic events within three months before treatment, connective tissue disease, severe acute infectious disease, stroke/neuro-surgery within six months before treatment, pregnancy, endocarditis, or ongoing anticoagulant treatment.[147] These patients would constitute the case group in the analysis of DNA hypermethylation as a biomarker for CRC detection. Furthermore, these patients were followed with regular intervals up to five years after inclusion. These patients would also be utilised in the assessment of circulating hypermethylated DNA as biomarkers for CRC prognosis.

During the same study period, blood samples from 143 patients referred for colonoscopy on suspicion of CRC were also included in the biobank. The colonoscopy revealed no sign of lower GI malignancy. These patients constituted the control group.

3.3. ETHICAL CONSIDERATIONS

All patients have provided written informed consent for the initial study. The North Denmark Region Committee on Health Research Ethics (N-20140064) and The Danish Data Protection Agency (2008-58-0028) have approved the conduction of the current study. The study was registered at ClinicalTrials.gov (NCT02928120).

3.4. BLOOD SAMPLING

All blood samples were obtained before colonoscopy in the control group and before any kind of treatment among CRC patients. Cubital venepuncture was used following the European Concerted Action on Thrombosis manual.[148] Sampling was conducted in EDTA tubes, centrifuged for 20 minutes at 4 °C and the plasma supernatant was subsequently stored at -80 °C. All blood samples were processed within two hours of collection.

3.5. HYPERMETHYLATION ANALYSIS

The detection of hypermethylated cytosines can be made using a variety of different molecular methods already described. The method used in this study is based on methylation specific real time PCR using a rapid bisulphite protocol.[141] The individual analytical steps are described below (a schematic version of the individual steps in the hypermethylation analysis is presented in Supplementary figure 1). The name and known function of each individual gene is provided in Supplementary table 1.

3.5.1. SAMPLE PREPARATION

We extracted sample nucleic acids from 350-1,000 µl of EDTA plasma using the easyMag™ (NucliSens® [bioMérieux SA, France]) according to the manufacturers' instructions. The extracted nucleic acids were subsequently eluted in 35 µl elution buffer (NucliSens® [bioMérieux SA, France]).

In order to quantitate the amount if DNA in each patient sample, we used five µl of the extracted DNA. We subsequently mixed the remaining 30 µl of purified nucleic acids with 60 µl of deamination solution, deaminated for ten minutes at 90 °C, following a purification step, and eluted in 25 µl 10 mM KOH.

3.5.2. PREAMPLIFICATION

The deaminated nucleic acids were then subject for a pre-amplification procedure in order to expand the limited amounts of DNA present in each patient sample, using the outer primers (Supplementary table 2 and 3).

The reaction buffer (25 µl) consisted of PCR buffer, 13 µM $MgCl_2$, 0.6 mM dNTP, 250 nM of each outer primer, 1.5 U Taq polymerase (MyTaq™ [Bioline®, Taunton, MA, USA]), and 0.3 U Cod Uracil-DNA Glycosylase (Cod UNG [ArcticZymes®, Tromsoe, Norway]). We distributed the first-round reaction mix to individual 200 µl PCR tubes, which were incubated for five minutes at 37 °C, followed by 95 °C for five minutes, and actively cooled to room temperature. Thereafter, we added 25 µl of purified deamination product to each tube, and performed the PCR reaction for 20 rounds (92 °C for 15 seconds, 55 °C for 30 seconds, and 72 °C for 30 seconds).

3.5.3. REAL-TIME PCR

The product from the pre-amplification procedure was used for the real-time PCR reaction using the inner primers and probes listed in Supplementary table 2 and 3.

For each individual reaction, we distributed 10 µl buffer containing 0.4 µM inner primers and probes in 30 individual wells in a 96 well PCR plate. We added 10 µl of the first round PCR product to 710 µl preincubated reaction mix (37 °C for five minutes and 95 °C for 10 minutes) containing PCR buffer, 250 µM dNTP, 10 µM $MgCl_2$, 8 U Taq polymerase (BIOTAQ™ [Bioline®, Taunton, MA, USA]), and 0.8 Uracil-DNA Glycosylase (Invitrogen®, Waltham, MA, USA). Twenty µl of reaction mix was then added to each of the 30 wells containing primers and probes. Real time PCR was conducted for 45 rounds (94 °C for 15 seconds, 55 °C for 30 seconds, and 72 °C for 30 seconds).

All PCR curves were analysed using the MxPro™, Mx3005P Quantitative PCR Software (Stratagene® [Agilent Technologies®, Santa Clara, CA, USA]) at a fluorescence threshold of 500 dR.

3.6. STATISTICAL ANALYSES

The evaluation of the individual hypermethylated DNA promoter regions as biomarkers for CRC detection and prognostication was done using two distinct analytical methods. The methods employed are described in detail according to each study. Each study is exploratory in nature.

In both studies, each hypermethylated biomarker was regarded as positive, if a signal was registered during PCR amplification. The amount of information lost due to this dichotomisation was evaluated using a stratification method, comparing the cycle threshold (Ct) value between CRC patients and controls. The median cell-free DNA levels were compared between CRC patients and healthy controls using the Wilcoxon-Mann-Whitney test, and between CRC stages using the Kruskal-Wallis test. Comparisons of the median number and range of cell-free hypermethylated promoter regions were made using the Wilcoxon-Mann-Whitney test.

STATA® V.13.1 (StataCorp. 2013. *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP) was used for all statistical analyses.

3.6.1. COLORECTAL CANCER DETECTION

To evaluate the utilisation of hypermethylated DNA fragments as biomarkers for CRC, we constructed a multivariable logistic regression model. The patients used for this analysis were patients with biopsy verified CRC and colonoscopy verified healthy controls. The stepwise model development and subsequent validation is presented below:

Outcome and predictor variables

The outcome variable was the presence of CRC. The potential predictor variables were each individual cell-free hypermethylated promoter region, CEA levels (>5 ng/ml), sex, and age. Each variable was handled as a dichotomous variable. All patients in the two groups were used in the analysis. The study was exploratory in nature, and as such, no sample size calculation was conducted.

Model development

Each of the potential hypermethylated predictor variables were analysed using univariate logistical regression. Each predictor variable reaching a significance level of $p \leq 0.3$ were included in the subsequent stepwise logistic regression analysis. The models were evaluated using receiver operating characteristics (ROC) curves and by calculating the area under the ROC curve (AUC). The model deemed most promising was then evaluated and each interaction between the potential predictor variables was estimated.

In order to ensure the correct selection of variables and assess the impact of separation issues in the model building process, we computed a penalised regression model using Firth's method with backwards selection.[149] Hosmer-Lemeshow's goodness of fit test was used to assess the final model.

Leave-pair-out cross validation was used to assess the inherent problem of model overfitting, and the optimism corrected AUC was presented to evaluate the predictive performance of the final model.

The predictive performance of the model was evaluated using all stage patients and subsequently evaluated in early stage CRC patients (stage I and II).

3.6.2. COLORECTAL CANCER PROGNOSIS

The study on the prognostic utility of cell-free hypermethylated DNA in plasma was divided into two parts: (i) a cross-sectional study on the correlation between the number of cell-free hypermethylated DNA promoter regions and the primary stage of CRC at the time of diagnosis, and (ii) a cohort study evaluating the impact of cell-free hypermethylated DNA on patient survival.

Outcome and predictor variables

The outcome variable for the cohort study was time from inclusion to death (or censoring). The potential prognostic variables were the 30 hypermethylated promoter regions previously described, pre-treatment CEA levels, along with sex and age at the time of diagnosis. The potential prognostic variables were handled as dichotomous variables (hypermethylated/un-methylated). CEA was considered positive if the level was above 5 ng/ml for non-smokers and above 10 ng/ml for smokers. Sex was handled as a categorical variable and as a continuous variable.

Survival analysis

As described in the introduction, stage is the universal predictor for survival among CRC patients. The number of hypermethylated promoter regions were therefore correlated to the primary stage of CRC. The non-parametric Wilcoxon-Mann-Whitney test was used to evaluate if the number differed between high and low stage or TNM classification (T3/4 vs T1/2, N1/2 vs N0, M1 vs M0).

Initially we screened each potential prognostic variable for their individual impact on patient survival using univariable cox-regression analysis. Using a significance level of $p<0.01$ we selected the potential prognostic variables to be analysed further in a multivariable Cox-regression model. In addition, the model incorporated the co-variables sex, age, pre-treatment CEA-levels, and primary CRC stage.

The Kaplan-Meier method was used to depict the association between cell-free hypermethylated promoter regions in plasma and patient survival.

Recurrence analysis

The risk of recurrence was evaluated using the data from inclusion on cell-free DNA hypermethylation. A competing risk analysis was conducted to evaluate if the number of hypermethylated DNA fragments measured at inclusion could predict disease recurrence. The Aalen-Johansen estimator was used to compute the cumulative incidence plot, and the pseudo-observation method was used to estimate the risk and the risk ratio (RR) of recurrence three years after inclusion.[150]

CHAPTER 4. RESULTS

This section will provide the reader with an overview of the results from the three studies performed. Moreover, the section evaluating the prognostic utility of cell-free hypermethylated DNA includes and their impact on CRC recurrence. The papers are appended as the final part of the thesis.

4.1. LITERATURE STUDY

To construct a prediction model for CRC detection and subsequent evaluation of the prognostic utility, a systematic literature search was conducted. This was to identify hypermethylated DNA fragments suggested relevant as biomarkers in the setting of CRC. The focus was putative hypermethylated promoter regions already evaluated in tumour remote media such as stool or blood samples.

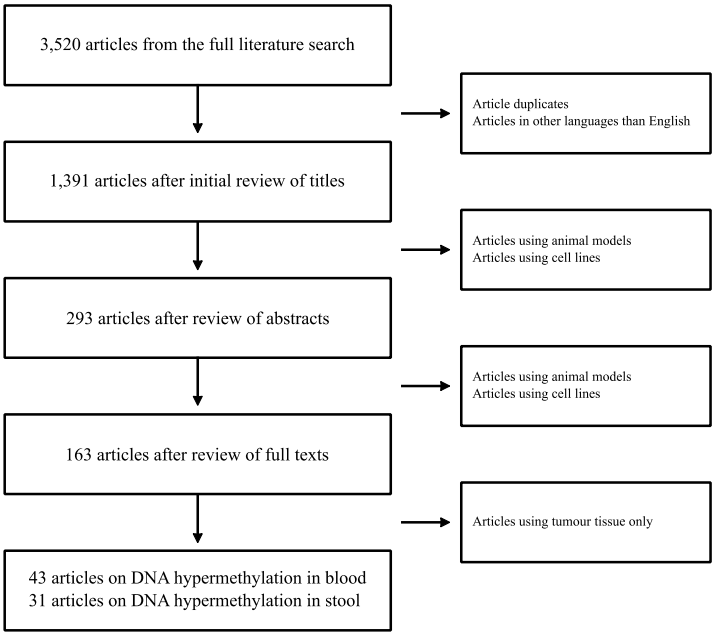


Figure 13 Flowchart for article selection

Search strategy in PubMed: ((((((((((tumor) OR stool sample) OR blood) OR serum) OR plasma)) AND ((((real time pcr) OR microarray analysis) OR genome sequencing) OR biopsy))) OR blood analysis)) AND Colorectal Neoplasms) AND ((DNA Methylation) OR CpG Islands)

We identified 3,520 potential articles in the different databases. Reviewing titles and removing duplicates left 1391 articles. Each abstract from the 1,391 articles was then reviewed leaving 74 articles; 43 analysing circulating DNA in blood samples and 31 analysing cell-free hypermethylated DNA in stool samples (Figure 13).

Most of the studies were characterised by few study participants. Only two large scale cross sectional studies have evaluated the performance of hypermethylated DNA fragments as biomarkers for CRC detection in stool and blood samples.[151,152] Tables 4 and 5 include the sensitivity and specificity for CRC, according to the individual gene promoter regions analysed.

Some biomarkers have been evaluated for their ability to detect early stage CRC and even precancerous lesion (*APC*, *NEUROG1*, *RASSF1A*, *RASSF2A*, *SDC2*, *SEPT9*, and *THBD*). [153–159] Albeit, the sensitivities for the individual markers for stage I CRC is rather modest (the sensitivity of *APC*, and *NEUROG1* was below 60% for stage I and II CRC), the biomarker *SDC2* reached a sensitivity of 92.3% for stage I CRC.[154] It therefore seems likely, that circulating hypermethylated DNA fragments could be used in blood based screening for CRC.

The biomarker, which has been subject to the most vigorous investigation, is *SEPT9*. *SEPT9* was identified from the assessment of >600 biomarker candidates.[160] Subsequently, the biomarker has been evaluated in numerous studies, with promising results.[155,161–167] However, the evaluation in a screening population only revealed a sensitivity of 48.2% at a specificity of 91.4%.[151] A recent study analysing a cohort of Danish CRC screening patients (published after the publication of our literature review) also revealed that the *SEPT9* assay is readily affected by the presence of arthritis and diabetes.[168] Moreover, age was associated with assay positivity, regardless of the presence of CRC. However, the biomarker still shows some promise as shown by a study by Jin et al. (2015).[169] They found that the *SEPT9* assay had a sensitivity of 74.8% at 87.4% specificity, compared to iFOBT with a sensitivity of 58.0% at 82.4% specificity. Interestingly, the CRCs "missed" by each test were not the same, and the authors suggest a combination of the two for future CRC screening.

SEPT9 has also been utilised as part of a biomarker panel (constituting *ALX4*, *SEPT9*, and *TMEFF2*) for CRC detection rendering a sensitivity of 81% at 90% specificity.[170] Moreover, the combination of *SEPT9* and *TAC1* had a sensitivity for stage I CRC of 73.1% at 92.3% specificity.[155] This favours the utilisation of multiple hypermethylated biomarkers for CRC detection.

Table 4 Blood-based hypermethylation biomarkers in colorectal cancer

Gene	*Specimen*	*Sensitivity, % (n)*		*Specificity, % (n)*		*Method*
ALX4	Serum	83.3%	(25/30)	70.0%	(21/30)[a]	qMSP
ALX4	Serum	46.6%	(21/45)	66.3%	(11/16)	qMSP
ALX4	Plasma	47.8%	(87/182)	93.5%	(159/170)	qMSP
APC	Serum	6.1%	(3/49)	100%	(41/41)[b]	qMSP
APC	Plasma	56.6%	(34/60)	86.0%	(86/100)[c]	qMSP
CDH1	Serum	17.6%	(3/17)	100%	(10/10)	MSP
CDH1	Plasma	60.0%	(36/60)	84.0%	(84/100)[c]	qMSP
CCND2	Plasma	96.7%	(29/30)	36.7%	(11/30)	MA
DAPK1	Plasma	50.0%	(30/60)	74.0%	(74/100)[c]	qMSP
FHIT	Plasma	50.0%	(30/60)	84.0%	(84/100)[c]	qMSP
HLTF	Serum	32.7%	(16/49)	92.7%	(38/41)[b]	qMSP
MLH1	Serum	42.9%	(21/49)	97.6%	(40/41)[b]	qMSP
NEUROG1	Serum	55.5%	(25/45)	81.3%	(13/16)	qMSP
CDKN2A	Serum	26.9%	(14/52)	100%	(44/44)[d]	MSP
CDKN2A	Serum	70.6%	(12/17)	100%	(10/10)	MSP
CDKN2A	Serum	6.6%	(14/211)	100%	(20/20)[e]	qMSP
RASSF1A	Serum	23.5%	(4/17)	100%	(10/10)	MSP
RASSF1A	Plasma	93.3%	(28/30)	53.3%	(16/30)	MA
RASSF1A	Serum	28.9%	(13/45)	90.0%	(81/90)[f]	MSP
RUNX3	Serum	64.7%	(11/17)	100%	(10/10)	MSP
RUNX3	Serum	41.5%	(27/65)	90.0%	(55/60)[g]	MSP
SDC2	Serum	87.0%	(114/131)	95.2%	(119/125)	qMSP
SEPT9	Plasma	72.0%	(90/125)	89.6%	(164/183)	MSP
SEPT9	Plasma	68.9%	(62/90)	89.0%	(138/155)	MSP
SEPT9	Serum	46.6%	(21/45)	81.3%	(13/16)	qMSP
SEPT9	Plasma	90.0%	(45/50)	88.3%	(83/94)	MSP
SEPT9	Plasma	95.6%	(88/92)	84.8%	(78/92)	MSP
SEPT9	Plasma	74.7%	(136/182)	96.0%	(164/170)	qMSP
SEPT9	Plasma	50.9%	(27/53)	91.4%	(1331/1457)[h]	qMSP
SEPT9	Plasma	74.8%	(101/135)	87.4%	(298/341)[i]	qMSP
SFRP2	Serum	66.9%	(113/169)	93.7%	(59/63)[j]	MSP
SMAD4	Plasma	52.0%	(31/60)	64.0%	(64/100)[k]	qMSP
TFPI2	Serum	18.1%	(39/215)	100%	(20/20)	qMSP
THBD	Plasma	70.7%	(53/75)	80.3%	(53/66)[l]	qMSP
TMEFF2	Plasma	70.9%	(129/182)	95.2%	(162/170)	qMSP
VIM	Serum	31.1%	(14/45)	62.5%	(10/16)	qMSP

Note: References for the individual studies are provided in the appended article. Unless otherwise specified, the control group was only specified as healthy. MSP = Methylation-specific PCR, qMSP = Quantitative methylation-specific PCR, MA = Microarray analysis.

[a]The control group included 15 patients with normal colons and 15 patients with low-grade inflammation or diverticulosis
[b]The control group included 41 patients with colonoscopy-verified healthy controls.
[c]The control group included 40 patients with adenomatous colorectal polyps of low-grade dysplasia.
[d]The control group included 37 patients with adenomatous colorectal polyps of low-grade dysplasia.
[e]The control group included 20 patients without malignancy and healthy controls.
[f]The control group included 30 patients with benign gastric disease, 30 patients with benign colorectal disease, and 30 healthy controls (not otherwise specified).
[g]The control group included 20 patients with benign oesophageal disease, 20 patients with benign gastric disease, and 20 patients with benign colorectal disease.
[h]The control group included 938 patients with no evidence of disease, 210 patients with non-advanced adenomas, and 315 patients with high-grade dysplasia, villous or ≥1 cm non-villous histology.
[i]The control group included 169 patients with adenomas, 81 with hyperplastic polyps, and 91 colonoscopy-verified healthy controls.
[j]The control group included 63 patients with benign colorectal adenomas.
[k]The control group included 40 patients with adenomatous colorectal polyps and 60 healthy controls.
[l]The control group included 66 patients with colonoscopy-verified healthy controls

Circulating hypermethylated DNA fragments have also yielded prognostic information. Some of the hypermethylation biomarkers are associated with advanced stage CRC (e.g. *ALX4*, *FBN2*, and *VIM*).[171–173] These individual hypermethylation biomarkers could therefore indicate the presence of micro or macro metastasis with poor prognosis. However, circulating hypermethylated DNA fragments are not solely surrogate markers for cancer stage. Wallner et al. (2006) showed that the presence of *HLTF* and/or *TMEFF1* hypermethylation inferred increased mortality (HR = 3.4, 95%CI[1.4; 8.1]), when adjusted for stage at the time of diagnosis. Further analysis of *HLTF* and *TMEFF1* by Philipp et al. (2012) showed that the biomarkers mainly provided prognostic information in stage IV CRC.[174] This shows that circulating hypermethylated DNA provides individual prognostic information, suggesting that they may be biomarkers for aggressive CRC.

Table 5 Stool-based hypermethylation biomarkers in colorectal cancer

Gene	*Sensitivity, % (n)*		*Specificity, % (n)*		*Method*
BMP3	100%	(9/9)	91.4%	(32/35)[a]	QuARTS
HIC1	42.3%	(11/26)	98.0%	(49/50)[b]	MSP
miR-34a	76.6%	(63/82)	95.0%	(38/40)	MSP
miR-34b/c	93.6%	(74/79)	100%	(40/40)	MSP
OSMR	37.7%	(26/69)	95.1%	(77/81)[c]	qMSP
NDRG4	60.7%	(17/28)	93.3%	(42/45)[d]	qMSP
PHACTR3	65.9%	(29/44)	100%	(30/30)	qMSP
SPG20	80.2%	(77/96)	100%	(30/30)[e]	MSP
SEPT9	20.0%	(7/35)	100%	(26/26)	PSQ
SFRP1	84.2%	(16/19)	85.7%	(12/14)	MSP
SFRP2	76.9%	(10/13)	76.9%	(10/13)	qMSP
SFRP2	94.2%	(49/52)	95.8%	(23/24)[f]	MSP
SFRP2	87.0%	(60/69)	93.3%	(28/30)[g]	qMSP
SFRP2	84.0%	(142/169)	54.0%	(34/63)[h]	MSP
TFPI2	75.8%	(50/66)	93.3%	(28/30)[i]	qMSP
TFPI2	68.3%	(41/60)	100%	(30/30)[j]	MSP
VIM	45.7%	(43/94)	90.0%	(178/198)	MSP
VIM	72.5%	(29/40)	86.9%	(106/122)[k]	MSP
VIM	38.3%	(23/60)	100%	(37/37)[l]	MSP
VIM	40.9%	(9/22)	94.7%	(36/38)	Methyl-beaming
VIM	54.5%	(18/35)	86.4%	(19/22)	PSQ

Note. References for the individual studies are provided in the appended article. Unless otherwise specified, the control group was only specified as healthy. MSP = Methylation-specific PCR; qMSP = Quantitative methylation-specific PCR; PSQ = Pyrosequencing; QuARTS = Quantitative allele-specific real-time target and signal amplification

[a]The control group included 35 controls with inflammatory bowel disease but without colorectal neoplasia.
[b]The control group included 32 controls with normal colonoscopies, 9 patients with hyperplastic polyps, and 9 patients with inflammatory bowel disease.
[c]The control group included 81 colonoscopy-verified healthy controls.
[d]The control group included 45 colonoscopy-verified healthy controls.
[e]The control group included 30 colonoscopy-verified healthy controls.
[f]The control group included 24 age-matched colonoscopy-verified healthy controls.
[g]The control group included 30 colonoscopy-verified healthy controls.
[h]The control group included 63 controls only specified as benign cases.
[i]The control group included 30 colonoscopy-verified healthy controls.
[j]The control group included 30 colonoscopy-verified healthy controls.
[k]The control group included 122 colonoscopy-verified healthy controls.
[l]The control group included 37 colonoscopy-verified healthy controls.

Studies analysing hypermethylated DNA from stool samples, have primarily focussed on candidate gene approaches. The presence of *BMP3*, *PHACTR3*, *SFRP2*, *SPG20*, *TFPI2*, *TMEFF2*, and *VIM* indicated early stage CRC or precancerous lesions.[175–180] Hypermethylated *VIM* was even commercialised (ColoSure®; Laboratory Corporation of America, Burlington, North Carolina, USA) , however it is not recommended for use by any agency and has largely been replaced by the FDA approved Cologuard® test (Exact Sciences Corporation, Madison, Wisconsin, USA). Cologuard® is a combined assay for aberrantly methylated *BMP3* and *NDRG4* promoter regions, *KRAS* mutation, β-actin, as well as an immunochemical assay for human haemoglobin. A large scale cross sectional study showed improved sensitivity for advanced precancerous lesions (42.4%) compared with iFOBT alone (23.8%).[152] However, the specificity of iFOBT was superior to the specificity achieved by the combined assay (96.4% and 89.8% respectively) leading to a false positive rate of the iFOBT of 3.6% and a false positive rate of the multitarget DNA test of 10.2%. The discovery of improved biomarkers for CRC detection could replace the markers currently implemented in the Cologuard® test.

Although the literature surrounding hypermethylated DNA fragments as biomarkers for CRC detection is growing rapidly, and *SEPT9* has been commercialised as a blood based CRC screening biomarker, there is some ambiguity in the current diagnostic accuracy studies. Presently, most of the studies conducted show a remarkable decrease in test performance from the initial development studies to the subsequent validation studies (e.g. *SEPT9*). This decrease is somewhat to be expected, however, it indicates that only utilising individual moleculat biomarkers, could lead to decreased test performance in CRC screening. The multitarget stool test (Cologuard®) and the results on the combination of *ALX4*, *SEPT9*, and *TMEFF2* as a blood based test, shows that biomarker panels of hypermethylated DNA fragments are more promising than individual hypermethylated biomarkers.

4.2. PATIENT POPULATION

The study population consisted of consecutive CRC patients referred for surgery. The 210 patients, from whom blood samples were available, were reviewed, leading to the exclusion of twelve patients: seven with benign disease or non-CRC, three without residual cancer after endoscopic resection, one patient who initially refused surgery, and one patient from whom informed consent could not be retrieved. Five additional patients were also excluded, because the reference gene (*MEST*) could not be amplified during PCR analysis (Supplementary figure 2).

Table 6 Patient characteristics

	Colorectal cancers		*Healthy controls*	
N	193		102	
Age, mean (SD)	67.5	(11.5)	64.7	(14.2)
Sex, n (%)				
Male	119	(61.7)	55	(53.9)
Female	74	(38.3)	47	(46.1)
Smoke status, n (%)				
Never smoker	68	(35.7)	31	(30.4)
Current smoker	77	(39.9)	28	(27.5)
Previous smoker	43	(22.3)	24	(23.5)
Unknown	5	(2.6)	19	(18.6)
CEA-levels				
≤ 5 ng/ml	141	(73.1)	91	(89.2)
> 5 ng/ml	52	(26.9)	11	(10.8)
Tumour, n (%)				
T1	3	(1.6)	-	-
T2	30	(15.5)	-	-
T3	120	(62.2)	-	-
T4	34	(17.6)	-	-
T-unknown	6	(3.1)	-	-
Node, n (%)				
N0	121	(62.7)	-	-
N1	38	(19.7)	-	-
N2	28	(14.5)	-	-
N-unknown	6	(3.1)	-	-
Metastasis, n (%)				
M0	159	(82.4)	-	-
M1	34	(17.6)	-	-

Note. The number (N) of patients in each patient group along with the mean age and standard deviations (SD) in each group. The number (n) and percentages (%) of patients according to sex, smoke status and carcinoembryonic antigen levels along with the colorectal cancer patients according to the tumour, node, and metastasis (TNM) classification system is also presented.

This left 193 CRC patients and 102 colonoscopy verified healthy controls. The remaining control group included 33 patients with resectable adenomas (all < 1 cm in size and without high grade dysplasia) (Table 6). All patients were used in the evaluation of hypermethylated DNA fragments as detection biomarkers for CRC. Only CRC patient samples were used in the evaluation of the prognostic properties of cell-free DNA hypermethylation.

4.3. COLORECTAL CANCER DETECTION

The 30 gene promoter regions chosen from the literature study were utilised as potential predictor variables for CRC detection. These hypermethylated promoter regions were evaluated in blood samples from the patients described above. The DNA concentrations in the plasma samples from the two patient populations were highly similar, with a median (range) of 4.10 ng/ml (0.31-52.19) in the control group and 4.00 ng/ml (0.26-132.58) in the CRC group (p = 0.982). The median number of hypermethylated DNA promoter regions (range) were four (1-11) and five (0-28) in the CRC group (p = 0.212). The correlation between DNA concentrations and the number of hypermethylated promoter regions is presented in Supplementary figure 3.

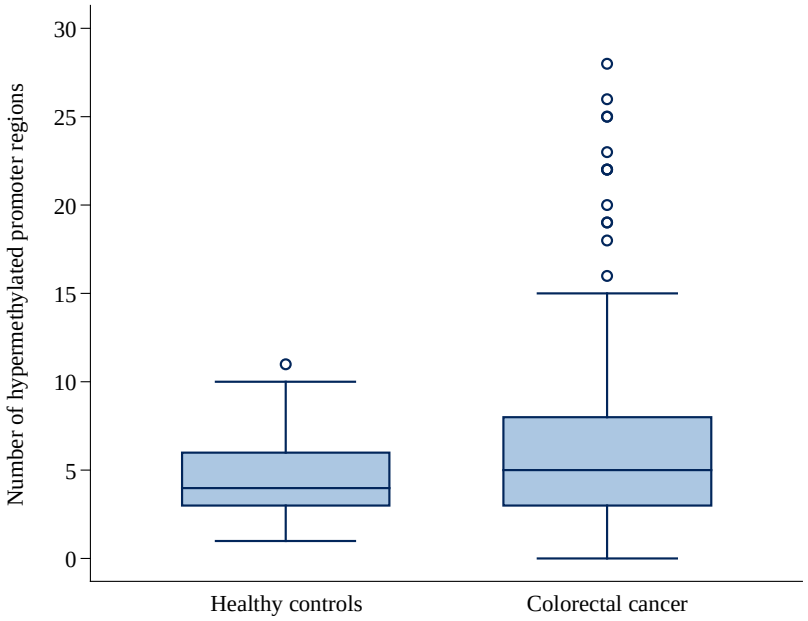


Figure 14 Number of hypermethylated promoter regions according to patient group

On an individual level, each of the hypermethylated DNA fragments had limited ability to distinguish CRC patients from healthy controls (Table 7). The marker showing most promise was *ALX4* with a sensitivity of 28.5% at 99.0% specificity. Other markers (e.g. *APC*, *NPTX2*, and *TAC1*) were more sensitive, but with lacking specificity.

Table 7 Promoter hypermethylation according to patient group

	Colorectal cancer(N=193)				*Healthy controls (N=102)*			
	N	*%*	*95% CI*		*N*	*%*	*95% CI*	
ALX4	55	28.5	22.2	35.4	1	1.0	0.0	5.3
APC	81	42.0	34.9	49.3	33	32.4	23.4	42.3
BMP3	55	28.5	22.2	35.4	11	10.8	5.5	18.5
BNC1	23	11.9	7.7	17.3	13	12.7	7.0	20.8
BRCA1	49	25.4	19.4	32.1	22	21.6	14.0	30.8
CDKN2A	18	9.3	5.6	14.3	4	3.9	1.1	9.7
HIC1	11	5.7	2.9	10.0	1	1.0	0.0	5.3
HLTF	22	11.4	7.3	16.7	4	3.9	1.1	9.7
MGMT	11	5.7	2.9	10.0	1	1.0	0.0	5.3
MLH1	87	45.1	37.9	52.4	44	43.1	33.4	53.3
NDRG4	18	9.3	5.6	14.3	0	0.0	0.0	3.6
NPTX2	135	69.9	62.9	76.3	60	58.8	48.6	68.5
NEUROG1	40	20.7	15.2	27.1	20	19.6	12.4	28.6
OSMR	22	11.4	7.3	16.7	7	6.9	2.8	13.6
PHACTR3	28	14.5	9.9	20.3	6	5.9	2.2	12.4
PPENK	20	10.4	6.4	15.6	4	3.9	1.1	9.7
RARB	49	25.4	19.4	32.1	71	69.6	59.7	78.3
RASSF1A	22	11.4	7.3	16.7	16	15.7	9.2	24.2
SDC2	47	24.4	18.5	31.0	6	5.9	2.2	12.4
SEPT9	47	24.4	18.5	31.0	5	4.9	1.6	11.1
SFRP1	42	21.8	16.2	28.3	7	6.9	2.8	13.6
SFRP2	39	20.2	14.8	26.6	18	17.6	10.8	26.4
SPG20	30	15.5	10.7	21.4	12	11.8	6.2	19.6
SST	58	30.1	23.7	37.1	32	31.4	22.5	41.3
TAC1	102	52.8	45.6	60.1	48	47.1	37.1	57.2
THBD	19	9.8	6.0	14.9	1	1.0	0.0	5.3
TFPI2	14	7.3	4.0	11.9	2	2.0	0.2	6.9
VIM	34	17.6	12.5	23.7	12	11.8	6.2	19.6
WIF1	19	9.8	6.0	14.9	4	3.9	1.1	9.7
WNT5A	12	6.2	3.3	10.6	5	4.9	1.6	11.1

Note. The number (N) of patients in each patient group along with the percentages (%) of patients according to each hypermethylated with corresponding 95% confidence intervals (95% CI)

The initial univariate screening left 19 of the potential predictor variables along with sex and age for the subsequent stepwise logistic regression. Through the stepwise selection process (Figure 15a) we considered Model 12 to be the most applicable due to a reduced model complexity, and because the model did not differ from the model

produced by penalised regression. Incorporation of CEA as a binary variable (> 5 ng/ml) did not lead to a difference in model selection, and CEA was therefore not incorporated in the model. The potential predictor variables were all evaluated for potential interactions, revealing an interaction between *RARB* and *VIM* ($p < 0.001$). However, the inclusion of this interaction term in the final model did not provide additional information. All models passed the Hosmer-Lemeshow test at the 0.05 significance level.

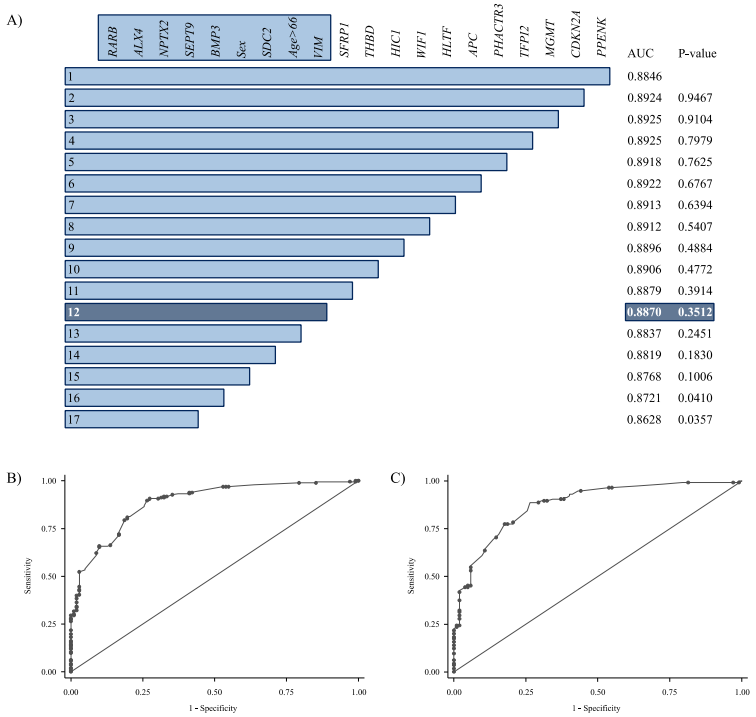


Figure 15 Stepwise selection process and model performance

A) Logistic regression with stepwise backwards selection. Potential predictor variables in the top row and model number in the left column. Area under the receiver operating characteristics (ROC) curve (AUC) are situated to the right. B) ROC curve for all stage CRC with a non-optimism corrected AUC of 0.8870. C) ROC curve for stage I and II CRC with a non-optimism corrected AUC of 0.8775.

Model 12 included seven hypermethylated promoter regions (*ALX4*, *BMP3*, *NPTX2*, *RARB*, *SDC2*, *SEPT9*, and *VIM*) along with sex and age. The model had the ability to distinguish CRC patients from colonoscopy verified healthy controls with an optimism corrected AUC of 0.860 (optimism = 0.027). Removal of 33 patients with adenomas from the control group did not alter the immediate efficacy of the prediction model (optimism corrected AUC = 0.863). Importantly, the model also had the ability to detect early stage CRC (stage I and II) with an optimism corrected AUC of

0.853 (optimism = 0.025). The model building process was done considering hypermethylation on a qualitative scale only. We therefore considered the Ct value of each individual DNA promoter region in order to assess the amount of information lost due to dichotomisation (Supplementary table 4). Inspection of the data suggests that there could be some difference in Ct value between CRC patients and healthy controls, suggesting that there was more of the individual markers present in the samples from the CRC patients. The effect of this difference in Ct value could, however, not be assessed in the current framework due to limited power.

4.4. COLORECTAL CANCER PROGNOSIS

In this part, we evaluated the same 30 cell-free hypermethylated DNA promoter regions as biomarkers for CRC stage and prognosis. This was in order to assess the added information provided by circulating hypermethylated DNA fragments to the current TNM staging system.

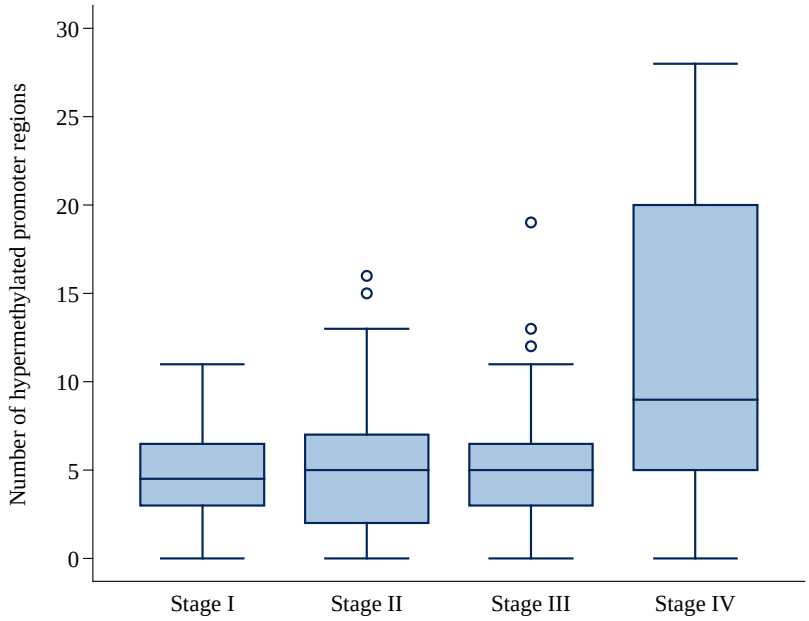


Figure 16 Number of hypermethylated promoter regions according to AJCC stage

The median cell-free DNA levels were similar between stages (3.73 ng/ml for stage I and 4.27 ng/ml for stage IV CRC) (p = 0.128).

The median number (range) of hypermethylated promoter regions in the CRC patients with distant metastasis was nine (0-28) compared to all other CRC patients with a median number of five (0-19) (p < 0.0001). However, the number of cell-free hypermethylated promoter regions were not associated with T- and N-classification.

Table 8 Promoter hypermethylation according to CRC stage

	AJCC stage 7th edition							
	I		*II*		*III*		*IV*	
	N	%	N	%	N	%	N	%
ALX4	5	2.6	19	9.8	13	6.7	18	9.3
APC	12	6.2	38	19.7	14	7.3	17	8.8
BMP3	5	2.6	27	14.0	7	3.6	16	8.3
BNC1	0	0.0	6	3.1	2	1.0	15	7.8
BRCA1	9	4.7	23	11.9	9	4.7	8	4.1
CDKN2A	4	2.1	6	3.1	1	0.5	7	3.6
HIC1	0	0.0	2	1.0	1	0.5	8	4.1
HLTF	1	0.5	6	3.1	6	3.1	9	4.7
MGMT	1	0.5	4	2.1	1	0.5	5	2.6
MLH1	12	6.2	42	21.8	18	9.3	15	7.8
NDRG4	1	0.5	5	2.6	4	2.1	8	4.1
NPTX2	18	9.3	56	29.0	34	17.6	27	14.0
NEUROG1	5	2.6	18	9.3	6	3.1	11	5.7
OSMR	1	0.5	7	3.6	2	1.0	12	6.2
PHACTR3	2	1.0	10	5.2	5	2.6	11	5.7
PPENK	3	1.6	4	2.1	3	1.6	10	5.2
RARB	3	1.6	20	10.4	11	5.7	15	7.8
RASSF1A	3	1.6	8	4.1	4	2.1	7	3.6
SDC2	4	2.1	15	7.8	9	4.7	19	9.8
SEPT9	3	1.6	18	9.3	7	3.6	19	9.8
SFRP1	3	1.6	9	4.7	9	4.7	21	10.9
SFRP2	3	1.6	8	4.1	10	5.2	18	9.3
SPG20	2	1.0	10	5.2	6	3.1	12	6.2
SST	7	3.6	21	10.9	11	5.7	19	9.8
TAC1	16	8.3	39	20.2	25	13.0	22	11.4
THBD	0	0.0	4	2.1	3	1.6	12	6.2
TFPI2	1	0.5	3	1.6	1	0.5	9	4.7
VIM	5	2.6	11	5.7	6	3.1	12	6.2
WIF1	2	1.0	3	1.6	2	1.0	12	6.2
WNT5A	0	0.0	4	2.1	2	1.0	6	3.1

Note. The number (N) and percentages (%) of colorectal cancer (CRC) patients (N=193) with positive amplification of hypermethylated promoter regions in plasma samples according to the American Joint Committee on Cancer (AJCC) staging system 7th Edition.

There was a marked increase in the frequency of promoter hypermethylation at all regions from stage I to stage II. However, for some of the promoter regions (e.g. *APC*, *MLH1*, *NPTX2*, and *TAC1*) there was an apparent decrease in methylation frequency from stage II to stage IV (Table 7). This indicates a large difference in methylation patterns between CRCs with metastasis, and tumour invasion.

There was an overall five-year survival of 62.2% (120/193) among the CRC patients. Patient survival was closely associated with the number of hypermethylated DNA fragments (Figure 17). The presence of more than four hypermethylated promoter regions was significantly associated with decreased overall survival ($p < 0.01$). The difference between having five to ten hypermethylated promoter regions in plasma and more than ten were only marginally different ($p = 0.09$).

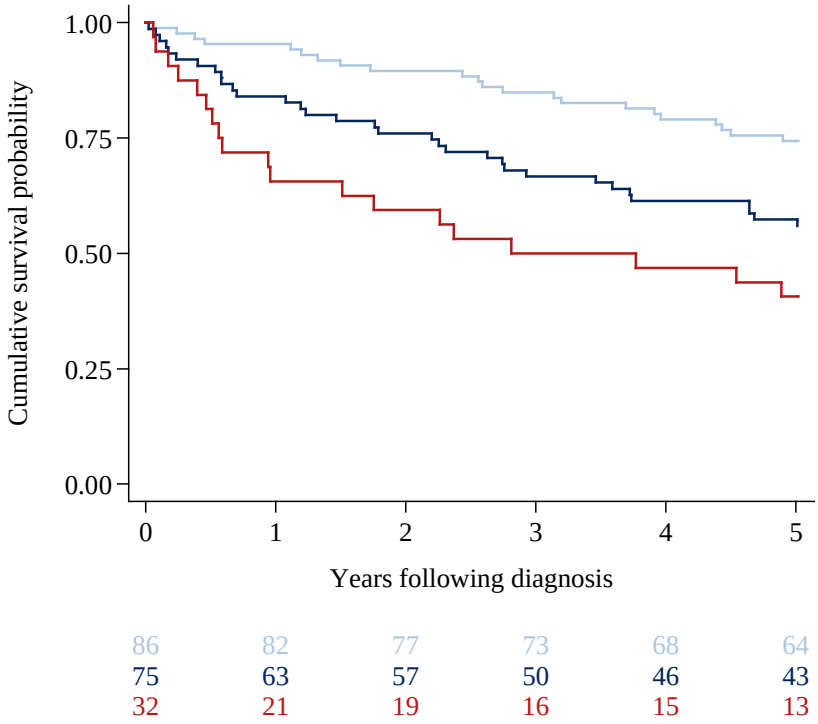


Figure 17 Survival according to hypermethylated DNA

The *light blue* line represents patients with 0-4, the *dark blue* line presents patients with 5-10, and the *red* line patients with more than 10 hypermethylated promoter regions measured in plasma. The number of patients at risk in each groups is presented below the graph.

Table 9 Hazard models

	Univariable Cox-regression				*Multivariable Cox-regression*			
	HR	*95% CI*		*P-value*	*HR*	*95% CI*		*P-value*
Sex	1.11	0.72	1.71	0.642	0.99	0.61	1.60	0.968
Age	1.00	1.00	1.00	0.005	1.00	1.00	1.00	<0.001
CEA	2.42	1.57	3.71	<0.001	1.50	0.89	2.52	0.126
Stage I	1.00				1.00			
Stage II	3.77	1.15	12.34	0.028	3.75	1.13	12.47	0.031
Stage III	4.99	1.48	16.88	0.010	5.09	1.45	17.78	0.011
Stage IV	22.36	6.76	73.92	<0.001	25.07	6.77	92.87	<0.001
ALX4	2.18	1.41	3.38	<0.001	1.50	0.86	2.63	0.157
BNC1	2.93	1.69	5.07	<0.001	1.44	0.64	3.26	0.377
HIC1	3.04	1.46	6.32	0.003	0.79	0.29	2.14	0.649
RARB	2.06	1.31	3.23	0.002	1.99	1.07	3.72	0.031
RASSF1A	2.75	1.57	4.81	<0.001	3.35	1.76	6.38	<0.001
SDC2	1.94	1.23	3.07	0.005	0.71	0.34	1.49	0.368
SEPT9	1.91	1.21	3.02	0.006	0.71	0.37	1.37	0.313
SFRP1	1.91	1.21	3.02	0.006	0.71	0.37	1.37	0.313
SFRP2	2.42	1.53	3.84	<0.001	0.98	0.46	2.06	0.955
SPG20	2.66	1.67	4.24	<0.001	1.73	0.85	3.51	0.131
TFPI2	2.67	1.38	5.19	0.004	0.92	0.33	2.52	0.863
THBD	2.95	1.68	5.18	<0.001	0.69	0.29	1.64	0.405
WIF1	3.26	1.83	5.83	<0.001	0.78	0.31	1.95	0.592
APC	1.19	0.77	1.82	0.440				
BMP3	1.45	0.92	2.29	0.106				
BRCA1	0.98	0.60	1.61	0.950				
CDKN2A	1.42	0.71	2.84	0.317				
HLTF	1.87	1.04	3.38	0.038				
MGMT	2.27	1.04	4.93	0.039				
MLH1	1.43	0.93	2.20	0.100				
NDRG4	1.54	0.80	2.99	0.197				
NPTX2	1.25	0.77	2.02	0.365				
NEUROG1	1.05	0.62	1.79	0.857				
OSMR	1.52	0.83	2.81	0.178				
PHACTR3	1.54	0.88	2.70	0.128				
PPENK	1.94	1.07	3.50	0.029				
SST	1.40	0.89	2.21	0.146				
TAC1	1.56	1.01	2.42	0.047				
VIM	1.82	1.10	3.00	0.020				
WNT5A	1.82	0.84	3.96	0.129				

Note: The univariable analysis of overall mortality using univariable Cox regression analysis. Variables reaching a significance level (p-values < 0.01) were analysed in the subsequent multivariable Cox regression analysis, adjusting for sex, age, CEA-levels and American Joint Committee on Cancer (AJCC) staging system. Individual hazard ratios (HR) with corresponding 95% confidence intervals (95% CI) and P-values. Carcinoembryonic antigen (CEA) was considered positive if the levels were above 5 mg/l for non-smokers, and if the levels were above 10 mg/l for smokers.

The majority of the hypermethylated DNA promoter regions inferred a poor prognosis. Thus, in order to limit the number of variables included in the multivariable Cox regression analysis we only included markers with a p-value below 0.01 (Table 8).

After adjustment for each of the biomarkers, stage was still the clearest predictor for adverse outcome. However, *RARB* and *RASSF1A* remained significantly associated with decreased overall survival, after adjusting for sex, age, pre-treatment CEA levels and stage. This decrease in overall survival remained, even after stratification for tumour localisation (colon versus rectum). In order to visualise the effect of *RARB* and *RASSF1A* on the survival of CRC patients, we computed the Kaplan Meier estimates for the positivity of the two markers (Figure 18).

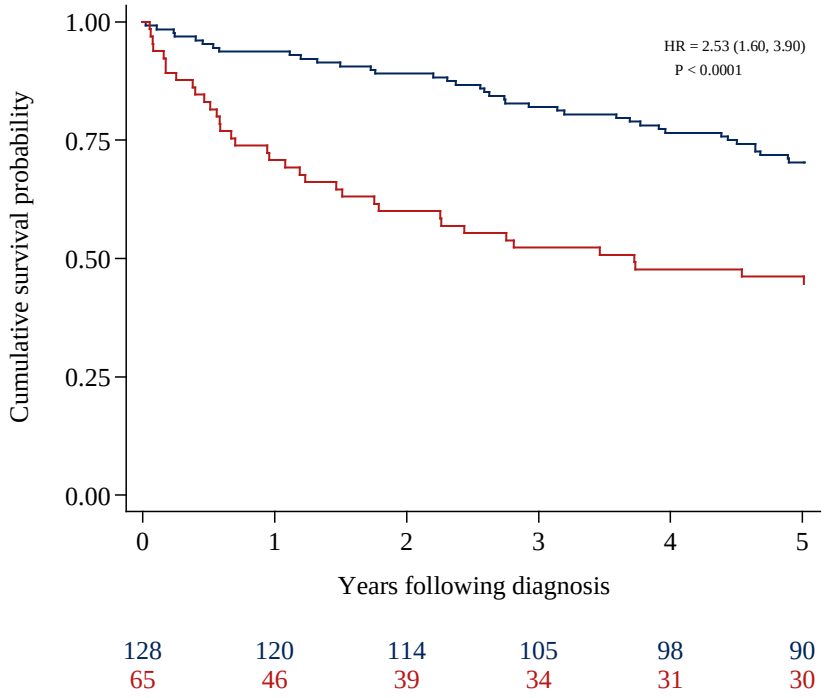


Figure 18 Survival according to *RARB* and *RASSF1A* hypermethylation

The *dark blue* line represents patients without hypermethylation of either promoter region and the *red* line represents patients with hypermethylated *RARB* and/or *RASSF1A*. The hazard ratio (HR) was computed using Cox regression with 95% confidence interval reported in brackets. The Log-rank test for equality of survivor functions was used to compute the p-value. The number of patients at risk in each groups is presented below the graph.

Of the 193 CRC patients, 144 received curatively intended surgical resections. Of these patients, 29 patients developed a local or metastatic recurrence. In order to assess if the number of circulating hypermethylated DNA promoter regions were associated with a risk of CRC recurrence, a cumulative incidence plot was computed for patients with 0-4 hypermethylated DNA promoter regions and patients with more than five (Figure 19).

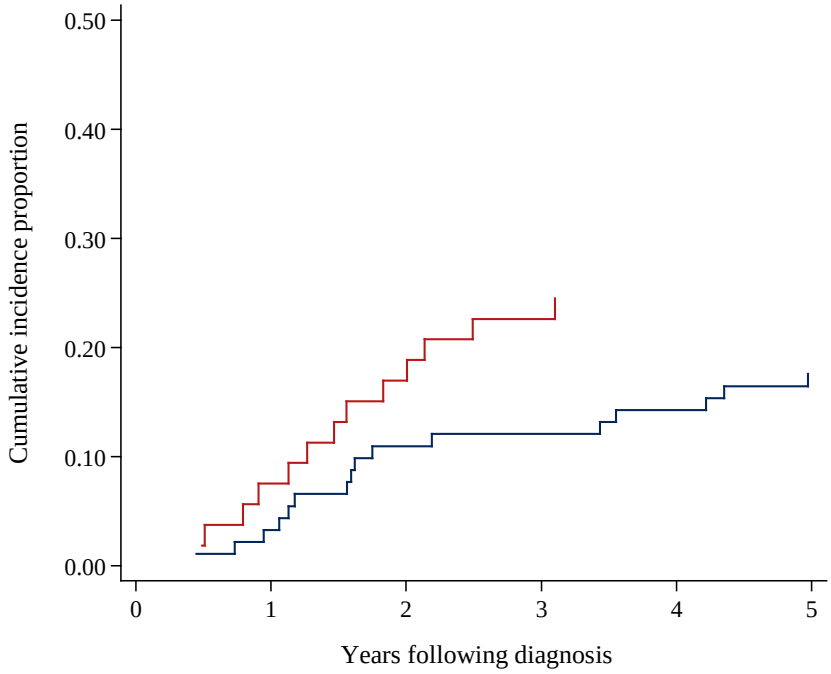


Figure 19 Risk of recurrence according to hypermethylated DNA

Cumulative incidence of recurrence when adjusting for death as a competing outcome (Aalen-Johansen estimator). The *dark blue* line represents patients with 0-4, and the *red* line represents patients with more than five hypermethylated promoter regions measured in plasma.

The risk of recurrence at three years after diagnosis was 12.8% in the group with a low number of circulating hypermethylated DNA promoter regions compared to 23.8% in the group with a high number. The nearly two-fold increased risk ratio at three years (RR = 1.86, 95%CI [0.88; 3.92]) was, however, not statistically significant (P = 0.101). Adjustment for stage did not alter the immediate conclusion (RR = 1.91, 95%CI [0.91; 4.04]). It was not possible to determine, if any of the individual markers inferred an increased risk of disease recurrence due to limited study power.

CHAPTER 5. DISCUSSION

Throughout the current thesis, a select few hypermethylated DNA promoter regions have been evaluated for their utility as blood based biomarkers for CRC detection and their association with the outcome of CRC patients. It appears, that circulating hypermethylated DNA measured in plasma (*ALX4*, *BMP3*, *NPTX2*, *RARB*, *SDC2*, *SEPT9*, and *VIM*) may be utilised as blood based biomarkers for CRC detection and some of them yield prognostic information (*RARB* and *RASSF1A*). Moreover, a high number of circulating hypermethylated DNA promoter regions seems to be related to decreased survival and a higher risk of disease recurrence.

5.1. LIMITATIONS

These results are indeed promising, however, before a blood based test can be utilised in the clinic, subsequent external validation must be conducted. Methodological limitations could lead to spurious associations, and these limitations will be elaborated below.

The first limitation encompasses the origin of the cell-free hypermethylated DNA fragments. The majority of circulating cell-free DNA originates from the turnover of normal cells.[181] We analysed our blood samples for the presence of putative biomarkers found through our initial literature study. We analysed these biomarkers in plasma samples only, and did not attempt to ensure their presence in paired patient tumours. This is an important caveat, because it makes us unable to ensure that the circulating hypermethylated DNA promoter regions are entirely tumour specific. In order to ensure the detection of tumour derived circulating DNA, we utilised the fact, that DNA from tumour cells is highly fragmented (mainly smaller than 150 bp) and designed our primers and probes accordingly (the largest amplified region was 152 bp (*WNT5A*)).[182] This does not ensure cancer specificity. However, regardless of origin, the performance of our biomarker panel for CRC remains promising (optimism corrected AUC = 0.860).

The second limitation is sample volume. The current thesis relies on the analysis of blood samples collected from a previous study.[147] One plasma sample was available for analysis for each patient, and in some instances, the sample volume was very low (350-1,000 µl plasma). As previously stated, the copy number of circulating tumour derived DNA is below 10 copies per 5,000 µl plasma.[123] Fortunately, the ability to detect our methylated reference gene (*MEST1*) was not affected by low plasma volumes (Supplementary figure 2), however, the limited amount of starting volume could lead to a decreased amount of cell-free tumour specific DNA for analysis. In

contrast, the next generation *SEPT9* test (Epi proColon® 2.0 [Epigenomics AG Corporation, Berlin, Germany]) requires a minimum of 3,500 µl plasma for DNA extraction.[162] Whether the low performance of *SEPT9* (sensitivity = 24.4%) reported in our study was merely a result of the limited amount of plasma remains unknown. Future evaluation of the developed biomarker panel should address this issue by isolating DNA from a larger volume of starting material.

The third limitation relates to the utilisation of stored blood samples. Prolonged storage of DNA may lead to substantial loss, with annual degradation rates of approximately 7-30%.[183,184] Another problem could be attributed to cell lysis due to prolonged processing time before storage, leading to a higher DNA yield originating from the intracellular space. This is an issue regardless of sample material (plasma/serum), but only markedly influencing the analysis of serum samples.[185] However, a recent study found that the degree of DNA methylation was not affected by prolonged processing (up to 72 hours) as long as samples were kept at room temperature.[186] In our study, we used EDTA plasma samples, and the processing never exceeded two hours. This suggests that the effect of processing issues should be limited. However, these considerations should be remembered in the future validation of the proposed biomarkers for CRC detection and to ensure that *RARB* and *RASSF1A* are indeed valid predictors for poor overall survival, regardless of stage in CRC patients.

The fourth limitation concerns the analytical method for hypermethylation detection.[141] Our method only allows detection of methylation at one site in the genome. Any discordance in primer/probe selection between the studies included in our review, and our current experiment could therefore lead to differences in marker performance. Moreover, we conducted our two-step PCR using methylation primers and probes in order to ensure the specificity of the assay. This could lead to false negative results, and a decrease in individual biomarker performance. However, we conducted this type of nested PCR in order to ensure the sensitivity of the assay for the limited amounts of hypermethylated DNA fragments detectable in plasma.

The fifth limitation revolves around the bisulphite conversion protocol. In our study we apply a rapid bisulphite protocol which should ensure complete conversion at an optimal DNA yield. This protocol was recently assessed to be suboptimal.[187] We have therefore compared kit with the second highest yield after bisulphite conversion (EZ DNA Methylation-Direct™; Zymo Research, Irvine, California, USA) and found that our test was more than comparable (Madsen P.H. [unpublished data]). The discrepancy between our results, and those presented by Ørntoft et al. (2017) may be due to subtle laboratory differences leading to decreased performance.[187] Future validation studies could employ more than one method to ensure optimal conversion and yield.

The sixth limitation surrounds multiple comparisons. In order to select biomarkers for CRC detection, we relied on a logistical selection process with backwards elimination. The elimination process is based on multiple significance tests, which can

lead to false positive associations. In the model development process, this phenomenon is coined overfitting. In order to estimate how much overfitting was present, we conducted *Leave-pair-out cross validation*.[188] This is in line with the current guidelines for studies surrounding predictive modelling.[189] However, the issue of false positive and false negative findings remains (with similar effects in the study on CRC prognosis), and external validation must be conducted, before any of the biomarkers can be utilised in the clinical setting.

The seventh and final limitation has previously been mentioned, and regards the lack of external validation. This is critical in order to properly assess the utility of the developed biomarker panel. However, inclusion of a validation cohort along with the analysis of plasma samples, was not possible in the predefined study period.

5.2. STRENGTHS

The main strength of the current study resides in the usage of a well-defined CRC cohort along with a select control population of colonoscopy verified healthy controls of a similar age and gender distribution. These CRC patients were included consecutively, and followed for up to five-years after diagnosis. We developed our biomarker panel using a rather large cohort compared to the studies already in the literature. Moreover, previous studies have often used less than optimal control groups for biomarker development, and validation of individual circulating hypermethylation biomarkers are generally lacking (one exception being *SEPT9*).[190]

Another strength is the use of multiple markers for CRC detection. Our study supports the results made by other groups, that a panel of hypermethylated biomarkers reaches a greater detection rate compared with individual biomarkers.[156,170] Even though the *SEPT9* test (Epi proColon®) has reached FDA approval as an alternative CRC screening test, it relies on the sole detection of *SEPT9* methylation in plasma. The implementation of multiple markers in CRC screening tests would lead to increased detection rates, as shown in the Cologuard® study.[152]

The current study relies on CRC detection using plasma based biomarkers. If these biomarkers are used in CRC screening, it could potentially increase adherence rates. In a study funded by Epigenomics AG, most patients non-adherent to an initial screening colonoscopy stated that they would prefer the *SEPT9* blood test (83%) over iFOBT (15%), with the remaining (2%) refusing any type of screening.[191] Whether an apparent preference for a blood based screening test will translate into improved adherence to the CRC screening programme in Denmark remains to be elucidated.

Lastly, our study is in full accordance with the current guidelines on predictive modelling.[189]

5.3. COMPARISON TO THE CURRENT LITERATURE

The results provided by the current study are generally comparable to the results found in the literature. Albeit the results are more modest.

Previous studies have found that the amount of circulating cell-free DNA is increased in cancer patients, with levels being increased approximately four times in CRC patients compared to healthy controls.[120,192] This is in contrast with the results from our analysis, where concentrations of cell-free DNA were similar (~4 ng/ml). However, it should be emphasised, that most studies on DNA concentrations in cancer patients have used blood donors as controls, which are not necessarily comparable to the otherwise healthy individuals present in a screening population.[193] Moreover, a significant increase in cell-free DNA is primarily found in serum samples, whereas we used plasma samples in our analysis. Whether the annual degradation rates of cell-free DNA has affected the CRC samples more compared to the controls is unknown, but seems highly speculative. An increase in DNA concentrations have more reliably been shown in lung and pancreatic cancer, with DNA concentrations being approximately four to five fold higher in patients compared to controls.[194,195] However, generally the overlap is rather large with reports of average concentrations of up to 44 ng/ml measured in plasma samples from healthy individuals.[196] This is in line with the findings in our study with ranges from 0.31 to 52.19 ng/ml for healthy controls and from 0.26 to 132.58 ng/ml for CRC patients. This shows some evidence against the utilisation of cell-free DNA levels as an immediate cancer detection biomarker - at least in the case of CRC.

5.3.1. DETECTION BIOMARKERS

The performance of each individual molecular marker for CRC in our biomarker panel is more modest than that of previous studies. None of the hypermethylated promoter regions analysed in our study could be used as individual biomarkers for CRC detection. With the exception of *SEPT9*, this is in line with results reported in the bulk of the present literature. The three most recent studies on *SEPT9* report sensitivities of 83.0-87.1% and specificities of 82.1-95.9%.[197–199] These results are indeed promising, however, the *SEPT9* assay has also been shown to be affected by age, stage of disease and benign disorders.[168] Moreover, the results from the PRESEPT study shows the real life screening scenario, and herein the assay sensitivity was only 48.2% for CRC and precancerous lesions.[151] As such, *SEPT9* cannot be recommended as an individual biomarker for CRC screening. However, as our results suggest, it may be utilised as part of a biomarker panel.

Another example of the limited use of individual CRC biomarkers is *NEUROG1*. *NEUROG1* methylations was identified and proved to be more specific than *ALX4* methylation with a sensitivity of 55.5% at 81.3% specificity.[158] In contrast, our analysis showed that not only was *ALX4* the superior of the two markers, *NEUROG1* was also equally methylated in CRC patients and healthy controls (~20%). This further suggests the use of multigene panels for CRC detection and promotes *ALX4* as a promising marker for CRC detection.

Previous reports on multigene panels for CRC detection have yielded promising results with *APC*, *MGMT*, *RASSF2A*, and *WIF1* providing a sensitivity of 86.5% at 92.1% specificity.[156] We evaluated three of these markers, and found that their individual performances were limited (sensitivities = 5.7-42.0% with specificities = 67.6-99.0%). The second largest study on combined hypermethylation biomarker panels for CRC detection evaluated *ALX4*, *SEPT9*, and *TMEFF2*.[170] The combination produced results comparable with our study (sensitivity of 80.7% at 90.0% specificity). Two of these markers were also implemented in our biomarker panel, however, the individual sensitivities were only 28.5% and 24.4%, respectively compared to 47.8% and 74.7% in the previous study.[170] This shows some of the difficulties in reproducing the results from putative biomarker studies. Some of this discrepancy could arise due to factors listed in the *limitations* section.

We evaluated *BMP3* and *NDRG4* due to the interesting findings in stool samples.[152] *BMP3* was implemented in the model, and while *NDRG4* was solely positive in CRC patients, it was only detected in 9.3% of CRC plasma samples. *NDRG4* therefore seems highly specific for CRC, without the necessary sensitivity required for an individual blood based CRC biomarker. This conclusion is supported by a recent report on *NDRG4* hypermethylation in 27% of plasma samples compared with 53% of stool samples in CRC patients.[200] The vast majority of colorectal tumours (like other cancer types) contain numerous cellular clones with only a subset of shared genetic or epigenetic alterations.[201] Differences in marker performance may be a result of this intratumour heterogeneity. A difference in gene expression at the invasive front compared with the central tumour area and the luminal part, has been proposed for β-catenin.[202] The molecular characteristics may therefore be different with regard to the invasion front (located near the invading blood vessels) and the luminal front. Another factor could be differences in the constitution of circulating DNA versus stool DNA. This makes the choice of media for molecular marker extraction critical. Future studies on biomarkers for CRC detection should keep this heterogeneity in mind when testing putative biomarker panels in stool or blood samples.

With the exception of *RARB*, all of the hypermethylated biomarkers implemented in the model for CRC detection were more frequent in the plasma of CRC patients compared to controls. *RARB* was found in 69.6% of healthy controls compared to 25.4% of CRC patients. Conversely, *RARB* methylation has been found in serum samples of 95.9% breast cancer patients compared to 0.0% in healthy women.[203] Moreover,

RARB has also been detected in rectal tumours with lymph node metastases.[204] In contrast, *RARB* was only hypermethylated in 1.6% of stage I CRC followed by a sharp increase to 20.7% for stage II and 7.8% for stage IV disease. This makes the utilisation of *RARB* as a biomarker for CRC enigmatic. On one hand, *RARB* hypermethylation is a biomarker for advanced tumours with some association to distant metastasis. On the other hand, *RARB* was more frequently hypermethylated in healthy controls above 66 years of age, and *RARB* methylation could likely just be a result of age related hypermethylation.[205] Whether our results concerning *RARB* are coincidental, requires further evaluation.

In the future, improvements and refinements in the area of EWAS on circulating cell-free DNA, as well as improvements in whole-genome bisulphite sequencing with minute amounts of starting material may lead to the detection of new biomarkers and possibly alter the performance of already existing CRC biomarkers.[206]

5.3.2. PROGNOSTIC BIOMARKERS

The data from our survival analysis quite clearly states, that increased CRC mortality is associated with an increase in circulating hypermethylated promoter regions.

Previous studies have shown that increased cell-free DNA levels were associated with decreased overall and recurrence free survival in CRC patients.[207] In our study, the levels of cell-free DNA did not differ between stage-groups, and were as such, not a good predictor for stage. We hypothesised, that the number of cell-free hypermethylated DNA promoter regions would be better suited as outcome predictors in CRC, and found it to be associated with metastatic disease. This is in line with previous results.[171–173] Moreover, a large number of hypermethylated DNA promoter regions were also a predictor for poor overall survival (possibly through the relation with stage IV disease). Univariate analysis also revealed that the majority of the analysed biomarkers were associated with poor overall survival in CRC patients.

Through our multivariable analyses, we discovered that *RARB* and *RASSF1A* were associated with poor survival in CRC patients, regardless of stage at the time of diagnosis. The presence of *RARB* and/or *RASSF1A* was associated with a two to three fold increase in overall mortality. Furthermore, an increased number of circulating hypermethylated DNA promoter regions were also associated with an increased risk of recurrence. This is in line with the studies on *HLTF* and *TMEFF2* with reports of a three to four fold increase in mortality for patients with either of the two markers, and a more than two-fold increase in the risk of recurrence among patients with circulating *HLTF*.[174,208] The following studies on *HLTF* and *TMEFF2*, however, showed that the impact on prognosis was limited to stage IV CRC patients.[209,210] We also

analysed the presence of *HLTF* methylation in our study, and while it was individually associated with poor survival (HR = 1.87, 95%CI [1.04; 3.38]) it was more often hypermethylated in patients with distant metastasis, and did not render prognostic information after adjustment for stage.

Recently, *SST* hypermethylation was shown to be a promising individual serum biomarker for overall survival in CRC.[211] We also evaluated circulating *SST* hypermethylation as a biomarker for CRC prognosis, and found no association. This difference could be attributed to some of the factors listed in the *limitations* section. The most important difference could be the use of plasma in the present work compared to the use of serum reported by others. The same group also evaluated *TAC1* hypermethylation and found no association with patient survival, which is in line with our results. This shows that there can be a marked difference in individual biomarker utilisation across studies and the comparison of results can be challenging.

In our experiment, we only assessed the methylation status of previously identified promoter regions in blood samples of CRC patients. However, since then, the number of tissue studies in CRC have increased rapidly. Recent tissue-studies have revealed new putative biomarkers, which provide prognostic information. *SHISA3* is transcriptionally repressed in colorectal tumours and adenomas, and *SHISA3* methylation has been shown to be a strong predictor for poor overall survival.[212] Other prognostic biomarkers identified through tissue studies may provide improved accuracy for the stratification of CRC patients according to prognosis.

The prognostic information provided by circulating hypermethylation biomarkers could aid in the choice of treatment for CRC patients. The presence of these individual biomarkers could imply the need for more intensive follow-up, and evaluation for the advent of local or metastatic recurrence, even after immediate curative resection. Whether our results on *RARB* and *RASSF1A* hypermethylation are reproducible warrants further validation in an independent sample of CRC patients.

5.4. FUTURE BIOMARKER RESEARCH

The search for new and improved biomarkers for CRC detection is ongoing. Currently the development of multi-biomarker hypermethylation assays are being readily developed and tested. Among these are the two-gene hypermethylation test (*BCAT1* and *IKZF1*) and the nucleosome test Nu.Q™ (Belgian Volition SPRL, Namur, Belgium).[213,214] The two-gene hypermethylation assay has recently been evaluated as an adjunct to the iFOBT screening programme increasing compliance rates in patients who were otherwise not interested in screening.[215] A planned large-scale Danish study conducted at Hvidovre Hospital, will include approximately 8000 patients in the evaluation of the Nu.Q™ test as a triage test after a positive iFOBT, selecting

patients for subsequent colonoscopy.[216] This could possibly lead to a decrease in the number of screening colonoscopies by ~25% (unpublished data). Whether the optimal utilisation of these blood based tests are as an up-front screening method (replacing the iFOBT) or as an adjunct to the current screening methods, in order to reduce the increasing burden of colonoscopies, remains to be established.

5.5. PERSPECTIVES

From the studies above it seems evident that there is both diagnostic and prognostic information rooted in these hypermethylated biomarkers.

For the validation of the current biomarker panel for CRC detection, we have included 102 patients with mild and moderate ulcerative colitis (November 2015 – July 2017) and 193 colonoscopy verified healthy controls from the screening population in Denmark (August 2016 – April 2017). We plan to compare the circulating DNA hypermethylation status of the genes identified through this thesis with approximately 100 CRC patients retrieved from the Danish Cancer Biobank. In order to raise the sensitivity of the assay, we plan to increase the amount of input plasma for DNA isolation by two-fold. Such a study will provide evidence for or against the efficacy of our proposed model for CRC detection. Moreover, this will prove, or disprove, the prognostic information gained from *RARB* and *RASSF1A* hypermethylation.

The current model for CRC detection using seven hypermethylated promoter regions is currently only on par with the current screening strategy using the iFOBT. A blood based biomarker panel should at least improve CRC detection, before it can replace the current method. Future studies will need to compare plasma-based assays and stool tests, in order to ensure improved performance by these tests. Furthermore, a thorough pipeline for biomarker evaluation should be employed to avoid waste and to increase reproducibility and comparison between laboratories.[217] Currently, the use of molecular biomarkers for CRC screening, have only provided increased sensitivity.[152] This enables the detection of more cancers, however, at the cost of decreased specificity. All patients who deliver a positive screening test have to undergo a subsequent colonoscopy for the final diagnosis, putting a large burden on local hospital departments. Before current screening methods (using iFOBT) eventually can be replaced by circulating or stool based hypermethylation biomarkers, any improvement in sensitivity must be accompanied by a specificity matching the specificity of the iFOBT.

CHAPTER 6. CONCLUSION

Circulating hypermethylated DNA fragments for cancer detection and prognosis have been thoroughly investigated for decades. However, none of these blood-based biomarkers are currently being recommended in the screening setting or as tools for CRC staging or prognostication.

In this thesis, the current knowledge surrounding cell-free hypermethylated DNA in stool and blood has been summarised. Even though several promising results have been proposed, most studies lack the reproducibility and performance justifying their immediate implementation in daily clinical practice.

We therefore set out to evaluate some of the putative biomarkers already evaluated in the literature, as markers for CRC detection. Individual hypermethylated promoter regions proved to be limited in their ability to detect CRC, however, a panel of seven gene promoter regions (*ALX4*, *BMP3*, *NPTX2*, *RARB*, *SDC2*, *SEPT9*, and *VIM*) along with sex and age were able to distinguish CRC patients from healthy controls with a promising performance (optimism corrected AUC of 0.853 for stage I and II CRC).

We also found that the number of hypermethylated promoter regions was associated with poor overall survival, and increased risk of recurrence. Moreover, we identified *RARB* and *RASSF1A* as predictors for increased mortality among CRC patients. These hypermethylated promoter regions could be used to identify patients who would benefit from a more intensive follow-up programme, or possibly adjuvant therapy.

These findings are only preliminary in nature and the methods utilised for the detection of cell-free promoter regions are rapidly evolving. This stresses the need for validation in external cohorts to ensure the exploratory results. Moreover, the utilisation of more input material for analysis could improve test sensitivity.

The years to come will elucidate if cell-free hypermethylated DNA fragments could be a valid and superior screening method compared with current stool based protocols.

LITERATURE

1 Engin O. *Colon Polyps and the Prevention of Colorectal Cancer*. Springer International Publishing, 2015 doi:10.1007/978-3-319-17993-3.

2 Fenger C. Histology of the anal canal. *Am J Surg Pathol* 1988; **12**: 41–55.

3 de Santa Barbara P, van den Brink GR, Roberts DJ. Development and differentiation of the intestinal epithelium. *Cell Mol Life Sci* 2003; **60**: 1322–1332.

4 Eckburg PB. Diversity of the Human Intestinal Microbial Flora. *Science* 2005; **308**: 1635–1638.

5 Chassaing B, Darfeuille–Michaud A. The Commensal Microbiota and Enteropathogens in the Pathogenesis of Inflammatory Bowel Diseases. *Gastroenterology* 2011; **140**: 1720–1728.e3.

6 Norman JM, Handley SA, Baldridge MT, Droit L, Liu CY, Keller BC *et al.* Disease-Specific Alterations in the Enteric Virome in Inflammatory Bowel Disease. *Cell* 2015; **160**: 447–460.

7 Bosca-Watts MM. Pathogenesis of Crohn's disease: Bug or no bug. *World J Gastrointest Pathophysiol* 2015; **6**: 1–12.

8 Louis P, Hold GL, Flint HJ. The gut microbiota, bacterial metabolites and colorectal cancer. *Nat Rev Microbiol* 2014; **12**: 661–672.

9 International Agency for Research on Cancer. GLOBOCAN 2012: Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012. [Online Material]. http://globocan.iarc.fr/ (accessed 15 Feb2016).

10 DCCG. Dansk Kolorektal Cancer Database Årsrapport. 2014.

11 Welch HG, Robertson DJ. Colorectal Cancer on the Decline — Why Screening Can't Explain It All. *N Engl J Med* 2016; **374**: 1605–1607.

12 Boyle T, Keegel T, Bull F, Heyworth J, Fritschi L. Physical activity and risks of proximal and distal colon cancers: A systematic review and meta-analysis. J. Natl. Cancer Inst. 2012; **104**: 1548–1561.

13 Lin KJ, Cheung WY, Lai JY-C, Giovannucci EL. The effect of estrogen vs.

combined estrogen-progestogen therapy on the risk of colorectal cancer. *Int J Cancer* 2012; **130**: 419–430.

14 Brenner H, Chang-Claude J, Seiler CM, Rickert A, Hoffmeister M. Protection From Colorectal Cancer After Colonoscopy. *Ann Intern Med* 2011; **154**: 22.

15 Elmunzer BJ, Hayward RA, Schoenfeld PS, Saini SD, Deshpande A, Waljee AK. Effect of Flexible Sigmoidoscopy-Based Screening on Incidence and Mortality of Colorectal Cancer: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *PLoS Med* 2012; **9**: e1001352.

16 Atkin WS, Edwards R, Kralj-Hans I, Wooldrage K, Hart AR, Northover JMA *et al.* Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. *Lancet* 2010; **375**: 1624–33.

17 Brenner H, Kloor M, Pox CP. Colorectal cancer. *Lancet* 2014; **383**: 1490–1502.

18 Aune D, Chan DSM, Lau R, Vieira R, Greenwood DC, Kampman E *et al.* Dietary fibre, whole grains, and risk of colorectal cancer: systematic review and dose-response meta-analysis of prospective studies. *BMJ* 2011; **343**: d6617–d6617.

19 Ma Y, Yang Y, Wang F, Zhang P, Shi C, Zou Y *et al.* Obesity and Risk of Colorectal Cancer: A Systematic Review of Prospective Studies. *PLoS One* 2013; **8**: e53916.

20 Chan DSM, Lau R, Aune D, Vieira R, Greenwood DC, Kampman E *et al.* Red and Processed Meat and Colorectal Cancer Incidence: Meta-Analysis of Prospective Studies. *PLoS One* 2011; **6**: e20456.

21 Fedirko V, Tramacere I, Bagnardi V, Rota M, Scotti L, Islami F *et al.* Alcohol drinking and colorectal cancer risk: an overall and dose-response meta-analysis of published studies. *Ann Oncol* 2011; **22**: 1958–1972.

22 Bosetti C, Rosato V, Gallus S, Cuzick J, La Vecchia C. Aspirin and cancer risk: a quantitative review to 2011. *Ann Oncol* 2012; **23**: 1403–1415.

23 Rothwell PM, Fowkes FGR, Belch JF, Ogawa H, Warlow CP, Meade TW. Effect of daily aspirin on long-term risk of death due to cancer: analysis of individual patient data from randomised trials. *Lancet* 2011; **377**: 31–41.

24 Burt R, Neklason DW. Genetic testing for inherited colon cancer. Gastroenterology. 2005; **128**: 1696–1716.

25 Jiao S, Peters U, Berndt S, Brenner H, Butterbach K, Caan BJ *et al.* Estimating the heritability of colorectal cancer. *Hum Mol Genet* 2014; **23**: 3898–3905.

26 Gryfe R. Inherited Colorectal Cancer Syndromes. *Clin Colon Rectal Surg* 2009; **22**: 198–208.

27 Siegel R, DeSantis C, Virgo K, Stein K, Mariotto A, Smith T *et al.* Cancer treatment and survivorship statistics, 2012. *CA Cancer J Clin* 2012; **62**: 220–241.

28 Iversen LH, Green A, Ingeholm P, Østerlind K, Gögenur I. Improved survival of colorectal cancer in Denmark during 2001–2012 – The efforts of several national initiatives. *Acta Oncol* 2016; **55**: 10–23.

29 Douaiher J, Ravipati A, Grams B, Chowdhury S, Alatise O, Are C. Colorectal cancer-global burden, trends, and geographical variations. *J Surg Oncol* 2017; **115**: 619–630.

30 Yabroff KR, Lund J, Kepka D, Mariotto A. Economic Burden of Cancer in the United States: Estimates, Projections, and Future Research. *Cancer Epidemiol Biomarkers Prev* 2011; **20**: 2006–2014.

31 Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RGS, Barzi A *et al.* Colorectal cancer statistics, 2017. *CA Cancer J Clin* 2017; **67**: 177–193.

32 Holme Ø, Bretthauer M, Fretheim A, Odgaard-Jensen J, Hoff G. Flexible sigmoidoscopy versus faecal occult blood testing for colorectal cancer screening in asymptomatic individuals. *Cochrane database Syst Rev* 2013; **9**: CD009259.

33 Lindebjerg J, Osler M, Bisgaard C. Colorectal cancers detected through screening are associated with lower stages and improved survival. *Dan Med J* 2014; **61**: A4758.

34 Kahi CJ, Rex DK, Imperiale TF. Screening, surveillance, and primary prevention for colorectal cancer: a review of the recent literature. *Gastroenterology* 2008; **135**: 380–99.

35 Faivre J, Dancourt V, Denis B, Dorval E, Piette C, Perrin P *et al.* Comparison between a guaiac and three immunochemical faecal occult blood tests in

screening for colorectal cancer. *Eur J Cancer* 2012; **48**: 2969–76.

36 DCCG. Dansk Kolorektal Cancer Database Årsrapport. 2013.

37 Holden DJ, Harris R, Porterfield DS, Jonas DE, Morgan LC, Reuland D *et al.* Enhancing the use and quality of colorectal cancer screening. *Evid Rep Technol Assess* 2010; 1–195.

38 Sewitch MJ, Fournier C, Ciampi A, Dyachenko A. Adherence to colorectal cancer screening guidelines in Canada. *BMC Gastroenterol* 2007; **7**: 39.

39 DTS. Dansk Tarmscreeningsdatabase Årsrapport. 2015.

40 Meester RGS, Doubeni CA, Zauber AG, Goede SL, Levin TR, Corley DA *et al.* Public health impact of achieving 80% colorectal cancer screening rates in the United States by 2018. *Cancer* 2015; **121**: 2281–2285.

41 Gupta AK, Pretlow TP, Schoen RE. Aberrant Crypt Foci: What We Know and What We Need to Know. *Clin Gastroenterol Hepatol* 2007; **5**: 526–533.

42 Muto T, Bussey HJ, Morson BC. The evolution of cancer of the colon and rectum. *Cancer* 1975; **36**: 2251–2270.

43 Eide TJ. Risk of colorectal cancer in adenoma-bearing individuals within a defined population. *Int J Cancer* 1986; **38**: 173–176.

44 Fleming M, Ravula S, Tatishchev SF, Wang HL. Colorectal carcinoma: Pathologic aspects. *J Gastrointest Oncol* 2012; **3**: 153–173.

45 Uraoka T, Hosoe N, Yahagi N. Colonoscopy: is it as effective as an advanced diagnostic tool for colorectal cancer screening? *Expert Rev Gastroenterol Hepatol* 2015; **9**: 129–132.

46 American Joint Committee on Cancer. Colon and rectum. In: *AJCC Cancer Staging Manual. 7th edition*. Springer: New York, 2010, pp 143–164.

47 Brierley JD, Gospodarowicz MK, Wittekind C. *The TNM Classification of Malignant Tumours, 8th Edition*. Wiley-Blackwell, 2016.

48 Puppa G, Sonzogni A, Colombari R, Pelosi G. TNM staging system of colorectal carcinoma: a critical appraisal of challenging issues. *Arch Pathol Lab Med* 2010; **134**: 837–852.

49 Lan Y-T, Yang S-H, Chang S-C, Liang W-Y, Li AF-Y, Wang H-S *et al.*

Analysis of the seventh edition of American Joint Committee on colon cancer staging. *Int J Colorectal Dis* 2012; **27**: 657–663.

50 Heald RJ. The 'Holy Plane' of rectal surgery. *J R Soc Med* 1988; **81**: 503–508.

51 Hohenberger W, Weber K, Matzel K, Papadopoulos T, Merkel S. Standardized surgery for colonic cancer: complete mesocolic excision and central ligation - technical notes and outcome. *Color Dis* 2009; **11**: 354–364.

52 West NP, Hohenberger W, Weber K, Perrakis A, Finan PJ, Quirke P. Complete Mesocolic Excision With Central Vascular Ligation Produces an Oncologically Superior Specimen Compared With Standard Surgery for Carcinoma of the Colon. *J Clin Oncol* 2010; **28**: 272–278.

53 Martel G, Crawford A, Barkun JS, Boushey RP, Ramsay CR, Fergusson DA. Expert Opinion on Laparoscopic Surgery for Colorectal Cancer Parallels Evidence from a Cumulative Meta-Analysis of Randomized Controlled Trials. *PLoS One* 2012; **7**: e35292.

54 Wu Z, Zhang S, Aung LHH, Ouyang J, Wei L. Lymph Node Harvested in Laparoscopic Versus Open Colorectal Cancer Approaches. *Surg Laparosc Endosc Percutan Tech* 2012; **22**: 5–11.

55 Ma Y, Yang Z, Qin H, Wang Y. A meta-analysis of laparoscopy compared with open colorectal resection for colorectal cancer. *Med Oncol* 2011; **28**: 925–933.

56 Bai H-L. Five-year long-term outcomes of laparoscopic surgery for colon cancer. *World J Gastroenterol* 2010; **16**: 4992–4997.

57 Kuhry E, Schwenk W, Gaupset R, Romild U, Bonjer HJ. Long-term results of laparoscopic colorectal cancer resection. In: Kuhry E (ed). *Cochrane Database of Systematic Reviews*. John Wiley & Sons, Ltd: Chichester, UK, 2008, p CD003432.

58 Bonjer HJ, Hop WCJ, Nelson H, Sargent DJ, Lacy AM, Castells A *et al.* Laparoscopically Assisted vs Open Colectomy for Colon Cancer. *Arch Surg* 2007; **142**: 298–303.

59 Schmoll H-J, Tabernero J, Maroun J, de Braud F, Price T, Van Cutsem E *et al.* Capecitabine Plus Oxaliplatin Compared With Fluorouracil/Folinic Acid As Adjuvant Therapy for Stage III Colon Cancer: Final Results of the NO16968 Randomized Controlled Phase III Trial. *J Clin Oncol* 2015; **33**:

3733–3740.

60 André T, Boni C, Navarro M, Tabernero J, Hickish T, Topham C *et al.* Improved Overall Survival With Oxaliplatin, Fluorouracil, and Leucovorin As Adjuvant Treatment in Stage II or III Colon Cancer in the MOSAIC Trial. *J Clin Oncol* 2009; **27**: 3109–3116.

61 Petersen SH, Harling H, Kirkeby LT, Wille-Jørgensen P, Mocellin S. Postoperative adjuvant chemotherapy in rectal cancer operated for cure. In: *Cochrane Database of Systematic Reviews*. 2012, p CD004078.

62 Van Cutsem E, Cervantes A, Adam R, Sobrero A, Van Krieken JH, Aderka D *et al.* ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. *Ann Oncol* 2016; **27**: 1386–1422.

63 Berber E, Pelley R, Siperstein AE. Predictors of Survival After Radiofrequency Thermal Ablation of Colorectal Cancer Metastases to the Liver: A Prospective Study. *J Clin Oncol* 2005; **23**: 1358–1364.

64 Padman S, Padbury R, Beeke C, Karapetis CS, Bishnoi S, Townsend AR *et al.* Liver only metastatic disease in patients with metastatic colorectal cancer: Impact of surgery and chemotherapy. *Acta Oncol* 2013; **52**: 1699–1706.

65 Labianca R, Nordlinger B, Beretta GD, Brouquet A, Cervantes A. Primary colon cancer: ESMO Clinical Practice Guidelines for diagnosis, adjuvant treatment and follow-up. *Ann Oncol* 2010; **21**: v70–v77.

66 Syk I, Wille Jörgensen P, Horváth-Puhó E, Høirup Petersen S, Hansdotter Andersson P, Martling A *et al.* The COLOFOL trial: study design and comparison of the study population with the source cancer population. *Clin Epidemiol* 2016; **8**: 15.

67 Tan E, Gouvas N, Nicholls RJ, Ziprin P, Xynos E, Tekkis PP. Diagnostic precision of carcinoembryonic antigen in the detection of recurrence of colorectal cancer. *Surg Oncol* 2009; **18**: 15–24.

68 Hardy PA, Zacharias H. Reappraisal of the Hansemann-Boveri hypothesis on the origin of tumors. *Cell Biol Int* 2005; **29**: 983–992.

69 Knudson AG. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A* 1971; **68**: 820–823.

70 Stehelin D, Varmus HE, Bishop JM, Vogt PK. DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian

DNA. *Nature* 1976; **260**: 170–173.

71 Knudson AG. Two genetic hits (more or less) to cancer. *Nat Rev Cancer* 2001; **1**: 157–162.

72 Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer Genome Landscapes. *Science* 2013; **339**: 1546–1558.

73 Kim S. New and emerging factors in tumorigenesis: an overview. *Cancer Manag Res* 2015; **7**: 225–239.

74 Vogelstein B, Fearon ER, Hamilton SR, Kern SE, Preisinger AC, Leppert M *et al.* Genetic Alterations during Colorectal-Tumor Development. *N Engl J Med* 1988; **319**: 525–532.

75 Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell* 1990; **61**: 759–67.

76 Seshagiri S, Stawiski EW, Durinck S, Modrusan Z, Storm EE, Conboy CB *et al.* Recurrent R-spondin fusions in colon cancer. *Nature* 2012; **488**: 660–664.

77 Muzny DM, Bainbridge MN, Chang K, Dinh HH, Drummond JA, Fowler G *et al.* Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 2012; **487**: 330–337.

78 Rowan AJ, Lamlum H, Ilyas M, Wheeler J, Straub J, Papadopoulou A *et al.* APC mutations in sporadic colorectal tumors: A mutational 'hotspot' and interdependence of the 'two hits'. *Proc Natl Acad Sci* 2000; **97**: 3352–3357.

79 Galiatsatos P, Foulkes WD. Familial Adenomatous Polyposis. *Am J Gastroenterol* 2006; **101**: 385–398.

80 Prior IA, Lewis PD, Mattos C. A Comprehensive Survey of Ras Mutations in Cancer. *Cancer Res* 2012; **72**: 2457–2467.

81 Barras D. BRAF Mutation in Colorectal Cancer: An Update. *Biomark Cancer* 2015; **7**: 9.

82 Pino MS, Chung DC. The Chromosomal Instability Pathway in Colon Cancer. *Gastroenterology* 2010; **138**: 2059–2072.

83 Vilar E, Gruber SB. Microsatellite instability in colorectal cancer—the stable evidence. *Nat Rev Clin Oncol* 2010; **7**: 153–162.

84 Lynch HT, Snyder CL, Shaw TG, Heinen CD, Hitchins MP. Milestones of Lynch syndrome: 1895–2015. *Nat Rev Cancer* 2015; **15**: 181–194.

85 Zaanan A, Cuilliere-Dartigues P, Guilloux A, Parc Y, Louvet C, de Gramont A *et al.* Impact of p53 expression and microsatellite instability on stage III colon cancer disease-free survival in patients treated by 5-fluorouracil and leucovorin with or without oxaliplatin. *Ann Oncol* 2010; **21**: 772–780.

86 Sargent DJ, Marsoni S, Monges G, Thibodeau SN, Labianca R, Hamilton SR *et al.* Defective Mismatch Repair As a Predictive Marker for Lack of Efficacy of Fluorouracil-Based Adjuvant Therapy in Colon Cancer. *J Clin Oncol* 2010; **28**: 3219–3226.

87 Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD *et al.* PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N Engl J Med* 2015; **372**: 2509–2520.

88 Allis CD, Jenuwein T. The molecular hallmarks of epigenetic control. *Nat Rev Genet* 2016; **17**: 487–500.

89 Baylin SB, Jones PA. Epigenetic Determinants of Cancer. *Cold Spring Harb Perspect Biol* 2016; **8**: a019505.

90 Smith ZD, Meissner A. DNA methylation: roles in mammalian development. *Nat Rev Genet* 2013; **14**: 204–220.

91 Wajed SA, Laird PW, DeMeester TR. DNA Methylation: An Alternative Pathway to Cancer. *Ann Surg* 2001; **234**: 10–20.

92 Timp W, Feinberg AP. Cancer as a dysregulated epigenome allowing cellular growth advantage at the expense of the host. *Nat Rev Cancer* 2013; **13**: 497–510.

93 Baylin SB, Jones PA. A decade of exploring the cancer epigenome — biological and translational implications. *Nat Rev Cancer* 2011; **11**: 726–734.

94 Prokhortchouk A, Sansom O, Selfridge J, Caballero IM, Salozhin S, Aithozhina D *et al.* Kaiso-Deficient Mice Show Resistance to Intestinal Cancer. *Mol Cell Biol* 2006; **26**: 199–208.

95 Lopes EC, Valls E, Figueroa ME, Mazur A, Meng F-G, Chiosis G *et al.* Kaiso Contributes to DNA Methylation-Dependent Silencing of Tumor Suppressor Genes in Colon Cancer Cell Lines. *Cancer Res* 2008; **68**: 7258–7263.

96 Lin P-C, Lin J-K, Lin C-H, Lin H-H, Yang S-H, Jiang J-K *et al.* Clinical Relevance of Plasma DNA Methylation in Colorectal Cancer Patients Identified by Using a Genome-Wide High-Resolution Array. *Ann Surg Oncol* 2015; **22**: 1419–1427.

97 Jair K-W. De novo CpG Island Methylation in Human Cancer Cells. *Cancer Res* 2006; **66**: 682–692.

98 De Marzo AM, Marchi VL, Yang ES, Veeraswamy R, Lin X, Nelson WG. Abnormal regulation of DNA methyltransferase expression during colorectal carcinogenesis. *Cancer Res* 1999; **59**: 3855–3860.

99 Clark SJ, Melki J. DNA methylation and gene silencing in cancer: which is the guilty party? *Oncogene* 2002; **21**: 5380–5387.

100 Kaz AM, Wong C-J, Dzieciatkowski S, Luo Y, Schoen RE, Grady WM. Patterns of DNA methylation in the normal colon vary by anatomical location, gender, and age. *Epigenetics* 2014; **9**: 492–502.

101 Spisák S, Kalmár A, Galamb O, Wichmann B, Sipos F, Péterfia B *et al.* Genome-Wide Screening of Genes Regulated by DNA Methylation in Colon Cancer Development. *PLoS One* 2012; **7**: e46215.

102 Figueiredo JC, Grau M V, Wallace K, Levine AJ, Shen L, Hamdan R *et al.* Global DNA Hypomethylation (LINE-1) in the Normal Colon and Lifestyle Characteristics and Dietary and Genetic Factors. *Cancer Epidemiol Biomarkers Prev* 2009; **18**: 1041–1049.

103 Sunami E, de Maat M, Vu A, Turner RR, Hoon DSB. LINE-1 Hypomethylation During Primary Colon Cancer Progression. *PLoS One* 2011; **6**: e18884.

104 Kim JC, Choi JS, Roh SA, Cho DH, Kim TW, Kim YS. Promoter methylation of specific genes is associated with the phenotype and progression of colorectal adenocarcinomas. *Ann Surg Oncol* 2010; **17**: 1767–1776.

105 Chan AO-O, Broaddus RR, Houlihan PS, Issa J-PJ, Hamilton SR, Rashid A. CpG island methylation in aberrant crypt foci of the colorectum. *Am J Pathol* 2002; **160**: 1823–1830.

106 Schnekenburger M, Diederich M. Epigenetics Offer New Horizons for Colorectal Cancer Prevention. *Curr Colorectal Cancer Rep* 2012; **8**: 66–81.

107 Moarii M, Reyal F, Vert J-P. Integrative DNA methylation and gene expression analysis to assess the universality of the CpG island methylator phenotype. *Hum Genomics* 2015; **9**: 26.

108 Weisenberger DJ, Siegmund KD, Campan M, Young J, Long TI, Faasse MA *et al.* CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nat Genet* 2006; **38**: 787–793.

109 Toyota M, Ahuja N, Ohe-Toyota M, Herman JG, Baylin SB, Issa J-PJ. CpG island methylator phenotype in colorectal cancer. *Proc Natl Acad Sci* 1999; **96**: 8681–8686.

110 Ogino S, Nosho K, Kirkner GJ, Kawasaki T, Meyerhardt JA, Loda M *et al.* CpG island methylator phenotype, microsatellite instability, BRAF mutation and clinical outcome in colon cancer. *Gut* 2009; **58**: 90–96.

111 Ogino S, Kawasaki T, Kirkner GJ, Kraft P, Loda M, Fuchs CS. Evaluation of Markers for CpG Island Methylator Phenotype (CIMP) in Colorectal Cancer by a Large Population-Based Sample. *J Mol Diagnostics* 2007; **9**: 305–314.

112 Anacleto C, Leopoldino AM, Rossi B, Soares FA, Lopes A, Rocha JCC *et al.* Colorectal Cancer 'Methylator Phenotype': Fact or Artifact? *Neoplasia* 2005; **7**: 331–335.

113 Jia M, Jansen L, Walter V, Tagscherer K, Roth W, Herpel E *et al.* No association of CpG island methylator phenotype and colorectal cancer survival: population-based study. *Br J Cancer* 2016; **115**: 1359–1366.

114 Hughes LAE, Melotte V, de Schrijver J, de Maat M, Smit VTHBM, Bovee JVMG *et al.* The CpG Island Methylator Phenotype: What's in a Name? *Cancer Res* 2013; **73**: 5858–5868.

115 Patai Á V, Valcz G, Hollósi P, Kalmár A, Péterfia B, Patai Á *et al.* Comprehensive DNA Methylation Analysis Reveals a Common Ten-Gene Methylation Signature in Colorectal Adenomas and Carcinomas. *PLoS One* 2015; **10**: e0133836.

116 Rashid A, Shen L, Morris JS, Issa J-PJ, Hamilton SR. CpG Island Methylation in Colorectal Adenomas. *Am J Pathol* 2001; **159**: 1129–1135.

117 Guinney J, Dienstmann R, Wang X, de Reyniès A, Schlicker A, Soneson C *et al.* The consensus molecular subtypes of colorectal cancer. *Nat Med* 2015;

21: 1350–1356.

118 Bramsen JB, Rasmussen MH, Ongen H, Mattesen TB, Ørntoft M-BW, Árnadóttir SS *et al.* Molecular-Subtype-Specific Biomarkers Improve Prediction of Prognosis in Colorectal Cancer. *Cell Rep* 2017; **19**: 1268–1280.

119 Mandel P, Metais P. Les acides nucléiques du plasma sanguin chez l'homme. *C R Seances Soc Biol Fil* 1948; **142**: 241–243.

120 Leon SA, Shapiro B, Sklaroff DM, Yaros MJ. Free DNA in the Serum of Cancer Patients and the Effect of Therapy. *Cancer Res* 1977; **37**: 646–650.

121 Schwarzenbach H, Hoon DSB, Pantel K. Cell-free nucleic acids as biomarkers in cancer patients. *Nat Rev Cancer* 2011; **11**: 426–437.

122 Swarup V, Rajeswari MR. Circulating (cell-free) nucleic acids - A promising, non-invasive tool for early detection of several human diseases. *FEBS Lett* 2007; **581**: 795–799.

123 Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N *et al.* Detection of Circulating Tumor DNA in Early- and Late-Stage Human Malignancies. *Sci Transl Med* 2014; **6**: 224ra24.

124 Olkhov-Mitsel E, Bapat B. Strategies for discovery and validation of methylated and hydroxymethylated DNA biomarkers. *Cancer Med* 2012; **1**: 237–260.

125 Rhoads A, Au KF. PacBio Sequencing and Its Applications. Genomics, Proteomics Bioinforma. 2015; **13**: 278–289.

126 Simpson JT, Workman RE, Zuzarte PC, David M, Dursi LJ, Timp W. Detecting DNA cytosine methylation using nanopore sequencing. *Nat Methods* 2017; **14**: 407–410.

127 Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc Natl Acad Sci* 1996; **93**: 9821–9826.

128 Liu L, Kron KJ, Pethe V V, Demetrashvili N, Nesbitt ME, Trachtenberg J *et al.* Association of tissue promoter methylation levels of APC, TGFβ2, HOXD3 and RASSF1A with prostate cancer progression. *Int J Cancer* 2011; **129**: 2454–2462.

129 Rawson JB, Manno M, Mrkonjic M, Daftary D, Dicks E, Buchanan DD *et*

al. Promoter methylation of Wnt antagonists DKK1 and SFRP1 is associated with opposing tumor subtypes in two large populations of colorectal cancer patients. *Carcinogenesis* 2011; **32**: 741–747.

130 Mrkonjic M, Roslin NM, Greenwood CM, Raptis S, Pollett A, Laird PW *et al.* Specific Variants in the MLH1 Gene Region May Drive DNA Methylation, Loss of Protein Expression, and MSI-H Colorectal Cancer. *PLoS One* 2010; **5**: e13314.

131 Matuschek C, Bölke E, Lammering G, Gerber PA, Peiper M, Budach W *et al.* Methylated APC and GSTP1 genes in serum DNA correlate with the presence of circulating blood tumor cells and are associated with a more aggressive and advanced breast cancer disease. *Eur J Med Res* 2010; **15**: 277–286.

132 Kron KJ, Liu L, Pethe V V, Demetrashvili N, Nesbitt ME, Trachtenberg J *et al.* DNA methylation of HOXD3 as a marker of prostate cancer progression. *Lab Investig* 2010; **90**: 1060–1067.

133 Bonanno C, Shehi E, Adlerstein D, Makrigiorgos GM. MS-FLAG, a Novel Real-Time Signal Generation Method for Methylation-Specific PCR. *Clin Chem* 2007; **53**: 2119–2127.

134 Cottrell SE. A real-time PCR assay for DNA-methylation using methylation-specific blockers. *Nucleic Acids Res* 2004; **32**: 10e–10.

135 Kristensen LS, Mikeska T, Krypuy M, Dobrovic A. Sensitive Melting Analysis after Real Time- Methylation Specific PCR (SMART-MSP): high-throughput and probe-free quantitative DNA methylation detection. *Nucleic Acids Res* 2008; **36**: e42–e42.

136 Wojdacz TK, Dobrovic A. Methylation-sensitive high resolution melting (MS-HRM): a new approach for sensitive and high-throughput assessment of methylation. *Nucleic Acids Res* 2007; **35**: e41–e41.

137 Worm J, Aggerholm A, Guldberg P. In-tube DNA methylation profiling by fluorescence melting curve analysis. *Clin Chem* 2001; **47**: 1183–1189.

138 Eads CA, Danenberg KD, Kawakami K, Saltz LB, Blake C, Shibata D *et al.* MethyLight: a high-throughput assay to measure DNA methylation. *Nucleic Acids Res* 2000; **28**: E32.

139 Warnecke PM, Stirzaker C, Song J, Grunau C, Melki JR, Clark SJ. Identification and resolution of artifacts in bisulfite sequencing. *Methods*

2002; **27**: 101–107.

140 Masser DR, Berg AS, Freeman WM. Focused, high accuracy 5-methylcytosine quantitation with base resolution by benchtop next-generation sequencing. *Epigenetics Chromatin* 2013; **6**: 33.

141 Pedersen I, Krarup H, Thorlacius-Ussing O, Madsen P. High recovery of cell-free methylated DNA based on a rapid bisulfite-treatment protocol. *BMC Mol Biol* 2012; **13**: 12.

142 Grunau C. Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Res* 2001; **29**: 65e–65.

143 Shiraishi M, Hayatsu H. High-speed conversion of cytosine to uracil in bisulfite genomic sequencing analysis of DNA methylation. *DNA Res* 2004; **11**: 409–415.

144 Hayatsu H, Negishi K, Shiraishi M. DNA methylation analysis: speedup of bisulfite-mediated deamination of cytosine in the genomic sequencing procedure. *Proc Japan Acad Ser B* 2004; **80**: 189–194.

145 Leontiou CA, Hadjidaniel MD, Mina P, Antoniou P, Ioannides M, Patsalis PC. Bisulfite conversion of DNA: Performance comparison of different kits and methylation quantitation of epigenetic biomarkers that have the potential to be used in non-invasive prenatal testing. *PLoS One* 2015; **10**: e0135058.

146 Huntriss JD, Hemmings KE, Hinkins M, Rutherford AJ, Sturmey RG, Elder K *et al.* Variable imprinting of the MEST gene in human preimplantation embryos. *Eur J Hum Genet* 2013; **21**: 40–47.

147 Tornby Stender M, Brøndum Frøkjær J, Sandie Hagedorn Nielsen T, Bjerregaard Larsen T, Lundbye-Christensen S, Elbrønd H *et al.* Combined use of clinical pre-test probability and D-dimer test in the diagnosis of preoperative deep venous thrombosis in colorectal cancer patients. *Thromb Haemost* 2008; **99**: 396–400.

148 Jespersen J, Bertina R, Haverkate F. *Laboratory Techniques in Thrombosis - A Manual. 2nd revised edition of ECAT Assay Procedures.* Kluwer Academic Publishers, 1999.

149 Wang X. Firth logistic regression for rare variant association tests. *Front Genet* 2014; **5**: 187.

150 Hansen SN, Andersen PK, Parner ET. Events per variable for risk differences

and relative risks using pseudo-observations. *Lifetime Data Anal* 2014; **20**: 584–598.

151 Church TR, Wandell M, Lofton-Day C, Mongin SJ, Burger M, Payne SR *et al.* Prospective evaluation of methylated SEPT9 in plasma for detection of asymptomatic colorectal cancer. *Gut* 2013; **63**: 317–325.

152 Imperiale TF, Ransohoff DF, Itzkowitz SH, Levin TR, Lavin P, Lidgard GP *et al.* Multitarget Stool DNA Testing for Colorectal-Cancer Screening. *N Engl J Med* 2014; **370**: 1287–1297.

153 Pack S-C, Kim H-R, Lim S-W, Kim H-Y, Ko J-Y, Lee K-S *et al.* Usefulness of plasma epigenetic changes of five major genes involved in the pathogenesis of colorectal cancer. *Int J Colorectal Dis* 2013; **28**: 139–147.

154 Oh T, Kim N, Moon Y, Kim MS, Hoehn BD, Park CH *et al.* Genome-Wide Identification and Validation of a Novel Methylation Biomarker, SDC2, for Blood-Based Detection of Colorectal Cancer. *J Mol Diagnostics* 2013; **15**: 498–507.

155 Liu Y, Tham CK, Ong SYK, Ho KS, Lim JF, Chew MH *et al.* Serum methylation levels of TAC1. SEPT9 and EYA4 as diagnostic markers for early colorectal cancers: a pilot study. *Biomarkers* 2013; **18**: 399–405.

156 Lee BB, Lee EJ, Jung EH, Chun H-K, Chang DK, Song SY *et al.* Aberrant Methylation of APC, MGMT, RASSF2A, and Wif-1 Genes in Plasma as a Biomarker for Early Detection of Colorectal Cancer. *Clin Cancer Res* 2009; **15**: 6185–6191.

157 Lange CPE, Campan M, Hinoue T, Schmitz RF, van der Meulen-de Jong AE, Slingerland H *et al.* Genome-Scale Discovery of DNA-Methylation Biomarkers for Blood-Based Detection of Colorectal Cancer. *PLoS One* 2012; **7**: e50266.

158 Herbst A, Rahmig K, Stieber P, Philipp A, Jung A, Ofner A *et al.* Methylation of NEUROG1 in serum is a sensitive marker for the detection of early colorectal cancer. *Am J Gastroenterol* 2011; **106**: 1110–1118.

159 Cassinotti E, Melson J, Liggett T, Melnikov A, Yi Q, Replogle C *et al.* DNA methylation patterns in blood of patients with colorectal cancer and adenomatous colorectal polyps. *Int J Cancer* 2012; **131**: 1153–1157.

160 Lofton-Day C, Model F, DeVos T, Tetzner R, Distler J, Schuster M *et al.* DNA Methylation Biomarkers for Blood-Based Colorectal Cancer

Screening. *Clin Chem* 2008; **54**: 414–423.

161 Ahlquist DA, Taylor WR, Mahoney DW, Zou H, Domanico M, Thibodeau SN *et al.* The Stool DNA Test Is More Accurate Than the Plasma Septin 9 Test in Detecting Colorectal Neoplasia. *Clin Gastroenterol Hepatol* 2012; **10**: 272–277.e1.

162 Tóth K, Sipos F, Kalmár A, Patai Á V, Wichmann B, Stoehr R *et al.* Detection of Methylated SEPT9 in Plasma Is a Reliable Screening Method for Both Left- and Right-Sided Colon Cancers. *PLoS One* 2012; **7**: e46000.

163 Warren JD, Xiong W, Bunker AM, Vaughn CP, Furtado L V, Roberts WL *et al.* Septin 9 methylated DNA is a sensitive and specific blood test for colorectal cancer. *BMC Med* 2011; **9**: 133.

164 Payne SR. From discovery to the clinic: the novel DNA methylation biomarker (m)SEPT9 for the detection of colorectal cancer in blood. *Epigenomics* 2010; **2**: 575–585.

165 Tänzer M, Balluff B, Distler J, Hale K, Leodolter A, Röcken C *et al.* Performance of epigenetic markers SEPT9 and ALX4 in plasma for detection of colorectal precancerous lesions. PLoS One. 2010; **5**: e9061.

166 DeVos T, Tetzner R, Model F, Weiss G, Schuster M, Distler J *et al.* Circulating methylated SEPT9 DNA in plasma is a biomarker for colorectal cancer. Clin. Chem. 2009; **55**: 1337–1346.

167 Grützmann R, Molnar B, Pilarsky C, Habermann JK, Schlag PM, Saeger HD *et al.* Sensitive Detection of Colorectal Cancer in Peripheral Blood by Septin 9 DNA Methylation Assay. *PLoS One* 2008; **3**: e3759.

168 Ørntoft M-BW, Nielsen HJ, Ørntoft TF, Andersen CL. Performance of the colorectal cancer screening marker Sept9 is influenced by age, diabetes and arthritis: a nested case–control study. *BMC Cancer* 2015; **15**: 819.

169 Jin P, Kang Q, Wang X, Yang L, Yu Y, Li N *et al.* Performance of a second-generation methylated SEPT9 test in detecting colorectal neoplasm. *J Gastroenterol Hepatol* 2015; **30**: 830–833.

170 He Q, Chen H-Y, Bai E-Q, Luo Y-X, Fu R-J, He Y-S *et al.* Development of a multiplex MethyLight assay for the detection of multigene methylation in human colorectal cancer. *Cancer Genet Cytogenet* 2010; **202**: 1–10.

171 Ebert MPA, Model F, Mooney S, Hale K, Lograsso J, Tonnes-Priddy L *et al.*

Aristaless-like Homeobox-4 Gene Methylation Is a Potential Marker for Colorectal Adenocarcinomas. *Gastroenterology* 2006; **131**: 1418–1430.

172 Hibi K, Mizukami H, Saito M, Kigawa G, Nemoto H, Sanada Y. FBN2 Methylation Is Detected in the Serum of Colorectal Cancer Patients with Hepatic Metastasis. *Anticancer Res* 2012; **32**: 4371–4374.

173 Shirahata A, Sakuraba K, Goto T, Saito M, Ishibashi K, Kigawa G *et al.* Detection of Vimentin (VIM) Methylation in the Serum of Colorectal Cancer Patients. *Anticancer Res* 2010; **30**: 5015–5018.

174 Wallner M, Herbst A, Behrens A, Crispin A, Stieber P, Göke B *et al.* Methylation of serum DNA is an independent prognostic marker in colorectal cancer. *Clin Cancer Res* 2006; **12**: 7347–52.

175 Kisiel JB, Yab TC, Hussain FTN, Taylor WR, Garrity-Park MM, Sandborn WJ *et al.* Stool DNA testing for the detection of colorectal neoplasia in patients with inflammatory bowel disease. *Aliment Pharmacol Ther* 2013; **37**: 546–554.

176 Bosch LJW, Oort FA, Neerincx M, Khalid-de Bakker CAJ, Terhaar Sive Droste JS, Melotte V *et al.* DNA methylation of phosphatase and actin regulator 3 detects colorectal cancer in stool and complements FIT. Cancer Prev. Res. 2012; **5**: 464–472.

177 Glöckner SC, Dhir M, Yi JM, McGarvey KE, Van Neste L, Louwagie J *et al.* Methylation of TFPI2 in stool DNA: a potential novel biomarker for the detection of colorectal cancer. *Cancer Res* 2009; **69**: 4691–4699.

178 Huang Z-H, Li L-H, Yang F, Wang J-F. Detection of aberrant methylation in fecal DNA as a molecular screening tool for colorectal cancer and precancerous lesions. World J. Gastroenterol. 2007; **13**: 950–954.

179 Zhang H, Song Y-C, Dang C-X. Detection of hypermethylated spastic paraplegia-20 in stool samples of patients with colorectal cancer. Int. J. Med. Sci. 2013; **10**: 230–234.

180 Javier Carmona F, Azuara D, Berenguer-Llergo A, Fernandez AF, Biondo S, de Oca J *et al.* DNA methylation biomarkers for noninvasive diagnosis of colorectal cancer. Cancer Prev. Res. 2013; **6**: 656–665.

181 Thierry AR, El Messaoudi S, Gahan PB, Anker P, Stroun M. Origins, structures, and functions of circulating DNA in oncology. *Cancer Metastasis Rev* 2016; **35**: 347–376.

182 Mouliere F, Robert B, Arnau Peyrotte E, Del Rio M, Ychou M, Molina F *et al.* High Fragmentation Characterizes Tumour-Derived Circulating DNA. *PLoS One* 2011; **6**: e23418.

183 Sozzi G, Roz L, Conte D, Mariani L, Andriani F, Verderio P *et al.* Effects of Prolonged Storage of Whole Plasma or Isolated Plasma DNA on the Results of Circulating DNA Quantification Assays. *J Natl Cancer Inst* 2005; **97**: 1848–1850.

184 Holdenrieder S, Von Pawel J, Nagel D, Stieber P. Long-term stability of circulating nucleosomes in serum. In: *Anticancer Research*. International Institute of Anticancer Research, 2010, pp 1613–1615.

185 Jung M. Changes in Concentration of DNA in Serum and Plasma during Storage of Blood Samples. *Clin Chem* 2003; **49**: 1028–1029.

186 Rasmussen L, Herzog M, Rømer E, Micallef J, Bulut O, Wilhelmsen M *et al.* Pre-analytical variables of circulating cell-free nucleosomes containing 5-methylcytosine DNA or histone modification H3K9Me3. *Scand J Clin Lab Invest* 2016; **76**: 448–453.

187 Worm Ørntoft M-B, Jensen SØ, Hansen TB, Bramsen JB, Andersen CL. Comparative analysis of 12 different kits for bisulfite conversion of circulating cell-free DNA. *Epigenetics* 2017; **2294**: 1–11.

188 Smith GCS, Seaman SR, Wood AM, Royston P, White IR. Correcting for Optimistic Prediction in Small Data Sets. *Am J Epidemiol* 2014; **180**: 318–324.

189 Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD Statement. *Br J Surg* 2015; **102**: 148–58.

190 Johnson DA, Barclay RL, Mergener K, Weiss G, König T, Beck J *et al.* Plasma Septin9 versus Fecal Immunochemical Testing for Colorectal Cancer Screening: A Prospective Multicenter Study. *PLoS One* 2014; **9**: e98238.

191 Adler A, Geiger S, Keil A, Bias H, Schatz P, DeVos T *et al.* Improving compliance to colorectal cancer screening using blood and stool based tests in patients refusing screening colonoscopy in Germany. *BMC Gastroenterol* 2014; **14**: 183.

192 Shapiro B, Chakrabarty M, Cohn EM, Leon SA. Determination of circulating DNA levels in patients with benign or malignant gastrointestinal disease.

Cancer 1983; **51**: 2116–2120.

193 Taback B, O'Day SJ, Hoon DSB. Quantification of circulating DNA in the plasma and serum of cancer patients. *Ann N Y Acad Sci* 2004; **1022**: 17–24.

194 Szpechcinski A, Chorostowska-Wynimko J, Struniawski R, Kupis W, Rudzinski P, Langfort R *et al.* Cell-free DNA levels in plasma of patients with non-small-cell lung cancer and inflammatory lung disease. *Br J Cancer* 2015; **113**: 476–483.

195 Henriksen SD, Madsen PH, Larsen AC, Johansen MB, Drewes AM, Pedersen IS *et al.* Cell-free DNA promoter hypermethylation in plasma as a diagnostic marker for pancreatic adenocarcinoma. *Clin Epigenetics* 2016; **8**: 117.

196 Fleischhacker M, Schmidt B. Circulating nucleic acids (CNAs) and cancer—A survey. *Biochim Biophys Acta - Rev Cancer* 2007; **1775**: 181–232.

197 Wu D, Zhou G, Jin P, Zhu J, Li S, Wu Q *et al.* Detection of Colorectal Cancer Using a Simplified SEPT9 Gene Methylation Assay Is a Reliable Method for Opportunistic Screening. *J Mol Diagnostics* 2016; **18**: 535–545.

198 Song L, Jia J, Yu H, Peng X, Xiao W, Gong Y *et al.* The performance of the mSEPT9 assay is influenced by algorithm, cancer stage and age, but not sex and cancer location. *J Cancer Res Clin Oncol* 2017; **143**: 1093–1101.

199 Song L, Peng X, Li Y, Xiao W, Jia J, Dong C *et al.* The SEPT9 gene methylation assay is capable of detecting colorectal adenoma in opportunistic screening. *Epigenomics* 2017; **9**: 599–610.

200 Melotte V, Yi JM, Lentjes MHFM, Smits KM, Van Neste L, Niessen HEC *et al.* Spectrin Repeat Containing Nuclear Envelope 1 and Forkhead Box Protein E1 Are Promising Markers for the Detection of Colorectal Cancer in Blood. *Cancer Prev Res* 2015; **8**: 157–164.

201 Melo FDSE, Vermeulen L, Fessler E, Medema JP. Cancer heterogeneity—a multifaceted view. *EMBO Rep* 2013; **14**: 686–695.

202 Brabletz T, Jung A, Hermann K, Günther K, Hohenberger W, Kirchner T. Nuclear Overexpression of the Oncoprotein β-Catenin in Colorectal Cancer is Localized Predominantly at the Invasion Front. *Pathol - Res Pract* 1998; **194**: 701–704.

203 Swellam M, Abdelmaksoud MDE, Sayed Mahmoud M, Ramadan A, Abdel-

Moneem W, Hefny MM. Aberrant methylation of APC and RARβ2 genes in breast cancer patients. *IUBMB Life* 2015; **67**: 61–68.

204 Leong KJ, Beggs A, James J, Morton DG, Matthews GM, Bach SP. Biomarker-based treatment selection in early-stage rectal cancer to promote organ preservation. *Br J Surg* 2014; **101**: 1299–1309.

205 Kristensen LS, Raynor MP, Candiloro I, Dobrovic A. Methylation profiling of normal individuals reveals mosaic promoter methylation of cancer-associated genes. *Oncotarget* 2012; **3**: 450–61.

206 Tanić M, Beck S. Epigenome-wide association studies for cancer biomarker discovery in circulating cell-free DNA: technical advances and challenges. *Curr Opin Genet Dev* 2017; **42**: 48–55.

207 Basnet S, Zhang Z-Y, Liao W-Q, Li S, Li P-S, Ge H-Y. The Prognostic Value of Circulating Cell-Free DNA in Colorectal Cancer: A Meta-Analysis. *J Cancer* 2016; **7**: 1105–1113.

208 Herbst A, Wallner M, Rahmig K, Stieber P, Crispin A, Lamerz R *et al.* Methylation of helicase-like transcription factor in serum of patients with colorectal cancer is an independent predictor of disease recurrence. *Eur J Gastroenterol Hepatol* 2009; **21**: 565–569.

209 Philipp AB, Stieber P, Nagel D, Neumann J, Spelsberg F, Jung A *et al.* Prognostic role of methylated free circulating DNA in colorectal cancer. Int. J. Cancer. 2012; **131**: 2308–2319.

210 Herbst A, Vdovin N, Gacesa S, Philipp A, Nagel D, Holdt LM *et al.* Methylated free-circulating HPP1 DNA is an early response marker in patients with metastatic colorectal cancer. *Int J Cancer* 2017; **140**: 2134–2144.

211 Liu Y, Chew MH, Tham CK, Tang CL, Ong SYK, Zhao Y. Methylation of serum SST gene is an independent prognostic marker in colorectal cancer. *Am J Cancer Res* 2016; **6**: 2098–2108.

212 Tsai M-H, Chen W-C, Yu S-L, Chen C-C, Jao T-M, Huang C-Y *et al.* DNA Hypermethylation of SHISA3 in Colorectal Cancer: An Independent Predictor of Poor Prognosis. *Ann Surg Oncol* 2015; **22**: 1481–1489.

213 Pedersen SK, Baker RT, McEvoy A, Murray DH, Thomas M, Molloy PL *et al.* A Two-Gene Blood Test for Methylated DNA Sensitive for Colorectal Cancer. *PLoS One* 2015; **10**: e0125041.

214 Rahier J-F, Druez A, Faugeras L, Martinet J-P, Géhénot M, Josseaux E *et al.* Circulating nucleosomes as new blood-based biomarkers for detection of colorectal cancer. *Clin Epigenetics* 2017; **9**: 53.

215 Symonds EL, Pedersen S, Cole SR, Massolino J, Byrne D, Guy J *et al.* Improving Participation in Colorectal Cancer Screening: a Randomised Controlled Trial of Sequential Offers of Faecal then Blood Based Non-Invasive Tests. *Asian Pacific J Cancer Prev* 2015; **16**: 8455–8460.

216 Rasmussen L, Wilhelmsen M, Christensen IJ, Andersen J, Jørgensen LN, Rasmussen M *et al.* Protocol Outlines for Parts 1 and 2 of the Prospective Endoscopy III Study for the Early Detection of Colorectal Cancer: Validation of a Concept Based on Blood Biomarkers. *JMIR Res Protoc* 2016; **5**: e182.

217 Ioannidis JPA, Bossuyt PMM. Waste, leaks, and failures in the biomarker pipeline. Clin. Chem. 2017; **63**: 963–972.

APPENDIX

Appendix A. Supplementary tables

Appendix B. Supplementary figures

Appendix C. Papers

Supplementary tables

Supplementary table 1 Gene names and known function

	Name	*Function*
ALX4	Aristaless-like homeobox 1	Skull and limb development
APC	Adenomatous polyposis	Cellular adhesion and β-catenin regulation
BMP3	Bone morphogenetic protein 3	Bone formation
BNC1	Basonuclin 1	Regulates proliferation and rRNA transcription
BRCA1	Breast cancer 1	DNA repair and gene transcription
CDKN2A	Cyclin dependent kinase inhibitor 2A	Cell-cycle regulation
HIC1	Hypermethylated in cancer 1	Gene transcription and cellular division
HLTF	Helicase like transcription factor	Gene transcription and cellular division
MGMT	O-6-methylguanine-DNA methyltransferase	DNA repair
MLH1	Mutl homolog 1	DNA repair
NDRG4	N-myc downstream-regulated gene 4	Cell-cycle progression and regulation of mitotic signals
NPTX2	Neuronal pentraxin 2	Synapse formation and non-apoptotic cell death
NEUROG1	Neurogenin 1	Neuronal differentiation
OSMR	Oncostatin M receptor	Cytokine receptor for cell signalling
PHACTR3	Phosphatase and actin regulator 3	Nuclear scaffolding of proliferating cells
PPENK	Preproenkephalin	Synaptic signalling
RARB	Retinoic acid receptor beta	Cellular signalling and retinoic acid binding
RASSF1A	Ras association domain family member 1A	Cell-cycle regulation and DNA repair
SDC2	Syndecan 2	Cellular adhesion, signalling, and cytoskeletal structure
SEPT9	Septin 9	Cell-cycle regulation and cytokinesis
SFRP1	Secreted frizzled related peptide 1	Modulator of Wnt signalling
SFRP2	Secreted frizzled related peptide 2	Modulator of Wnt signalling
SPG20	Spastic paraplegia 20	Regulates endosomal traffic and mitochondria function
SST	Somatostatin	Inhibitory hormone and regulator of endocrine system
TAC1	Tachykinin precursor 1	Neurotransmitting and vasodilation
THBD	Thrombomodulin	Inhibition of haemostasis
TFPI2	Tissue factor pathway inhibitor 2	Inhibition of haemostasis
VIM	Vimentin	Cell shape and integrity maintenance
WIF1	Wnt inhibitory factor 1	Cell fate regulation in oncogenesis and embryogenesis
WNT5A	Wnt family member 5A	Cell fate regulation in oncogenesis and embryogenesis

Note. Gene names and known function have been cross-matched with the RefSeq database (http://www.ncbi.nlm.nih.gov/refseq/)

Supplementary table 2 Characteristics of the reference gene (*MEST*) primers and probes

	Outer primers		*Inner primers*		*Probes*	*Coordinates*	*Accession no.*
MEST U	(+) GGTTTTAAAAGTTCGGTGTTTATT (−) CCIAACAACTACAACCACTCC	(130)	TGTTGTGGTAATTAGTATATTTT CAACCACTCCAACATACACTACA	(83)	CGCGAGTAGTTGTGTTTTGTTCGCG	130492052 – 130492181	NC_000007.14
MEST M	(+) GCGATGGGTTTGTGCGC (−) GAAAAACCGATTACGCATACG	(130)	CGACGTTTTAGTTTCGAGTC CGCTTCCTAAAACCAAAAATTCTCG	(86)	CGATCGGTGGTCGGGTTCGATCG	130486225 – 130486355	NC_000007.14

Note. The primer and probe sequences for the methylation specific polymerase chain reaction with the individual amplicon sizes represented as number of base pairs in brackets to the right of the outer/inner primers respectively. The NCBI accession number (no.) and the remapping coordinates of the amplified promoter regions are also presented. (+) Forward primer, (−) Reverse primer. (*U*) indicates the un-methylated version of *MEST*, (*M*) indicates the methylated version of *MEST*.

Supplementary table 3 Characteristics of gene specific primers and probes

		Outer primers		*Inner primers*		*Probes*	*Coordinates*	*Accession no.*
ALX4	(+)	CGTTTTCGTTCGTCGTTTGC	(114)	TTTTTCGGAGGCGATAAGTTC	(85)	CGCGATTGTCGGTCGTCGTTAAAGTATCGCG	44309869 – 44309983	NC_000011.10
	(−)	CGAACCCGACTCTTAACG		CGAACCCGACTCTTAACG				
APC	(+)	ATTGCGGAGTGCGGGTC	(98)	AGTCGGGGTCGGGAAGC	(91)	CGCGATCGTTGGATGCGGAATCGCG	112737725 – 112737823	NC_000005.10
	(−)	AATCGACGAACTCCCGACG		AATCGACGAACTCCCGACG				
BMP3	(+)	TAGCGTTGGAGTGGAGAC	(114)	AGTGGAGACGGCGTTC	(96)	CGTCGAGCGGGTGAGGTTCGCGTATCGACG	81031015 – 81031129	NC_000004.12
	(−)	CCAACCCCACTTACTACG		CTTACTACGCTAACCCAACG				
BNC1	(+)	GTAGGTAGTTAGTTGGTTTTC	(102)	GTAGGTAGTTAGTTGGTTTTC	(83)	CGCGATCGTATTTACGGGAGTCGGAGTTTGATCGCG	83284403 – 83284505	NC_000015.10
	(−)	GCGAAAATTCTCTATACG		GAAACAAACGACCCGAAACG				
BRCA1	(+)	GTTTTTTGGTTTTCGTGGTAAC	(114)	TCGTGGTAACGGAAAAGCGCG	(83)	CGATCGGCGGCGTGAGCGTACG	43125420 – 43125441	NC_000017.11
	(−)	AAACCCACAACCTATCCCCCG		CCGTCCAAAAAATCTCAACG				
CDKN2A	(+)	TGTTCGGAGTTAATAGTATTTTTTTC	(133)	TTTCGAGTATTCGTTTATAGC	(111)	CGACGTGAAAGATATCGCGGTACTTCG	21974925 – 21975058	NC_000009.12
	(−)	TTTCTTCCTCCGATACTAACG		TTTCTTCCTCCGATACTAACG				
HIC1	(+)	GATATAACGTTTTTTTCGCGTC	(142)	TTCGGTTTTCGCGTTTTGTTC	(91)	CGCGACGGTCGTCGTTCGGGTTCGCG	2056054 – 2056196	NC_000017.11
	(−)	ATACCGCCCTAACGCCG		CGAAAACTATCAACCCTCG				
HLTF	(+)	TTTGATTGTTGTAGAAGGAGAC	(121)	TAGAAGGAGACGGCGTC	(88)	CGCGATCGATTGGATTCGCGGCGAGATCGCG	149086415 – 149086536	NC_000003.12
	(−)	CCGCAACCCTAAAACCG		AAAACAAATTCCGAACGCCG				
MGMT	(+)	GATATGTTGGGATAGTTCGC	(129)	GATATGTTGGGATAGTTCGC	(119)	CGCGATCGTATCGTTTGCGATTTATCGCG	129467213 – 129467342	NC_000010.11
	(−)	AAAAAACTCCGCACTTCCG		GCACTCTTCCGAAACGAAACG				
MLH1	(+)	TGGTTTTTTGGCGTTAAAATGTC	(123)	TGGTTTTTTGGCGTTAAAATGTC	(94)	CGCGATCTCGTCCAACCGCCGAATATCGCG	36993529 – 36993652	NC_000003.12
	(−)	CATCTCTTTAATAACATTAACTAACCG		AAATAACTTCCCCCGCCG				
NDRG4	(+)	GTATTTTAGTCGCGTAGAAGGC	(111)	TTACGCGCGAGGGGATC	(88)	CGCGATCGCGGTTCGTTCGGGATTAGTTGATCDCG	58463524 – 58463635	NC_000016.10
	(−)	AATTTAACGAATATAAACGCTCG		AATTTAACGAATATAAACGCTCG				
NPTX2	(+)	TTCGGTAGGTTAGAGTGTC	(94)	AGGTTAGAGTGTCGAGTAGC	(80)	CGCGATCGGTGCGGTTGTGAGACGGTGATCGCG	98617274 – 98617368	NC_000007.14
	(−)	CTATCGTCTCGAAAATCGCG		TCGAAAATCGCGTACACCG				
NEUROG1	(+)	GTTTATACGAGTTGATTGATC	(92)	GTTGATTTGATGTCGGC	(63)	CGCGATGCCCGACCGATCTCCTAAATCGCG	135535860 – 135535952	NC_000005.10
	(−)	CTTAACCTAACCTCCTCG		CTCGCCTACAAAAACCACG				
OSMR	(+)	TCGGTGTCGTTTTTGTTTAGC	(147)	TCGGTGTCGTTTTTGTTTAGC	(97)	CGCGATCTTCGGACGGCGTTCGGATCGCG	38846295 – 38846442	NC_000005.10
	(−)	AAACGACTCGACACTATACG		CCCACCAAATAAACAAACCG				
PHACTR3	(+)	GAGGTTAGGAGAGGTTAAGC	(98)	AGGTTAAGGATTCGTTC	(82)	CGCGATCCAACAACGCTAAAACCGCTGATCGCG	59577157 – 59577255	NC_000020.11
	(−)	AACGAAAAATATAACCGCCG		GCCGATTCCCCGCG				
PPENK	(+)	CGCGTTATTTCGGGAATC	(133)	AGGCGATTTGAGTCGTTTTTAC	(112)	CGCGATCAAAGTTGTCGGTCGGGAGGATCGCG	56446032 – 56446165	NC_000008.11
	(−)	GACAACCTCAACAAAAATCG		GACAACCTCAACAAAAATCG				
RARB	(+)	AGTAGGGTTTGTTTGGGTATC	(127)	GGGTATCGTCGGGGTAGATTC	(113)	CGCGACGAATACGTTCCGAATCGCG	25428388 – 25428515	NC_000003.12
	(−)	TCGACCAATCCAACCGAAACG		TCGACCAATCCAACCGAAACG				
RASSF1A	(+)	GGGAGGCGTTGAAGTC	(115)	GGGAGGCGTTGAAGTC	(76)	CGCGATTCGTTCGGTTCGCTCGCG	50340782 – 50340897	NC_000003.12
	(−)	AATAAACTCAAACTCCCCCG		GTACTTCGCTAACTTTAAACG				
SDC2	(+)	GCGTAGGAGGAGGAAGC	(102)	GTTTCGAGTTCGAGTTTTC	(66)	CGCGATCCGTGAGTCGTAATCGTTGCGGTGATCGCG	96494093 – 96494195	NC_000008.11
	(−)	CAACCCGCGCACACG		CGCACACGAATCCG				
SEPT9	(+)	GTTTAGTATTTATTTTCGAAGTTC	(118)	GTTTAGTATTTATTTTCGAAGTTC	(93)	CGACGTATTTAGTTGCGCGTTGATCGACGTCG	77373442 – 77373560	NC_000017.11
	(−)	GCCGAAAACGCTTCCTCG		CCTCCGCGACGACCG				
SFRP1	(+)	GAGGCGATTGGTTTTCGC	(149)	GGAGTTGATTGGTTGCGC	(90)	CGCGATGGTTCGGTCGTAATCGCG	41309435 – 41309584	NC_000008.11
	(−)	CGCGACACTAACTCCG		CGCGACACTAACTCCG				
SFRP2	(+)	GTTTTTCGGAGTTGCGCGC	(124)	GTTTTTCGGAGTTGCGCGC	(94)	CGACGTTTGTAGCGTTTCGTTCGCGTTGTTACGTCG	153788922 – 153789046	NC_000004.12
	(−)	CTCTTCGCTAAATACGACTCG		CCGAAAAACTAACAACCGACG				
SPG20	(+)	GAACGTTTTGGTTGTTAC	(142)	GAACGTTTTGGTTGTTAC	(123)	CGCGATCCGCTACGCTCGCCGAAATCGCG	36346256 – 36346398	NC_000013.11
	(−)	CTCGAAAACTCCCTACG		ACCGCCGCGACAACG				
SST	(+)	AGAGTATATAAGTCGTTTAGGAGC	(65)	AGTCGAGATGTTGTTTTGTC	(81)	CGATCGACCAACGCGCACTAACGATCG	187670234 – 187670315	NC_000003.12
	(−)	CCAAAACCAAAACGATAAACAACG		CCAAAACCAAAACGATAAACAACG				
TAC1	(+)	TAAGGAGGTTGGGATAAATATC	(83)	ATCGTAAGGTATTGAGTAGGC	(64)	CGATCGATCCGAACGCGCTCTCGATCG	97732034 – 97732126	NC_000007.14
	(−)	TCTCGATAACTACCGCCG		TCTCGATAACTACCGCCG				
THBD	(+)	CGTCGTAGTAATTTATTTGTTATC	(119)	CGTCGTAGTAATTTATTTGTTATC	(99)	CGCGATCGTTCGGATTTCGGCGTTTGATCGCG	23050377 – 23050496	NC_000020.11
	(−)	TAAACACGCACTAACCAACG		CCAAACCCCATCTCATCG				
TFPI2	(+)	TATTTTTTAGGTTTCGTTTCGGC	(86)	TATTTTTTAGGTTTCGTTTCGGC	(72)	CGCGATCGTCGGTCGGACGTTCGTTGATCGCG	93890745 – 93890831	NC_000007.14
	(−)	CGACTTTCTACTCCAAACG		AAACGACCCGAATACCCG				
VIM	(+)	GAGGTTTTCGCGTTAGAGAC	(143)	ATATTTATCGCGTTTTCGTTC	(102)	CGACGTGTTCGCGTTATCGTCGTCGACGTCG	17229296 – 17229439	NC_000010.11
	(−)	ACGAACCTAATAAACATAACTACG		ACGAACCTAATAAACATAACTACG				
WIF1	(+)	CGACGCGTTTAGTC	(109)	GTTGAGGGAGTTGTAGC	(77)	CGCGATCGGTTGCGTAGGTGCGGGATCGCG	65121211 – 65121320	NC_000012.12
	(−)	TACCGAAAAAACTCCTCG		TACCGAAAAAACTCCTCG				
WNT5A	(+)	CGTGGAATAGTTGTTTGC	(152)	CGTGGAATAGTTGTTTGC	(134)	CGCGATCAACCTAATCGAAACGCAACTAAAGATCGCG	55487159 – 55487311	NC_000003.12
	(−)	CGAACCTAAACTCCCG		TTAAAACAAAACTAAAATACG				

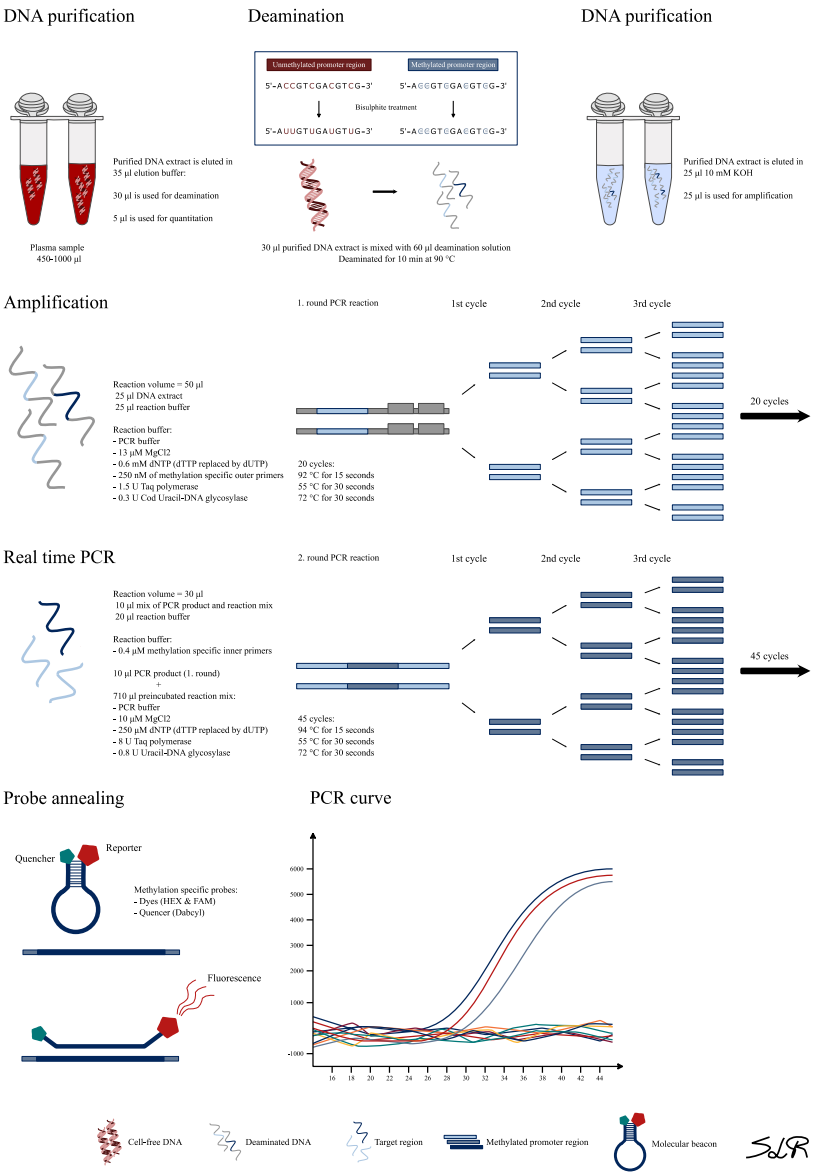
Note. The primer and probe sequences for the methylation specific polymerase chain reaction with the individual amplicon sizes represented as number of base pairs in brackets to the right of the outer/inner primers respectively. The NCBI accession number (no.) and the remapping coordinates of the amplified promoter regions are also presented. (+) Forward primer, (−) Reverse primer.

Supplementary table 4 Patients and controls according to cycle threshold

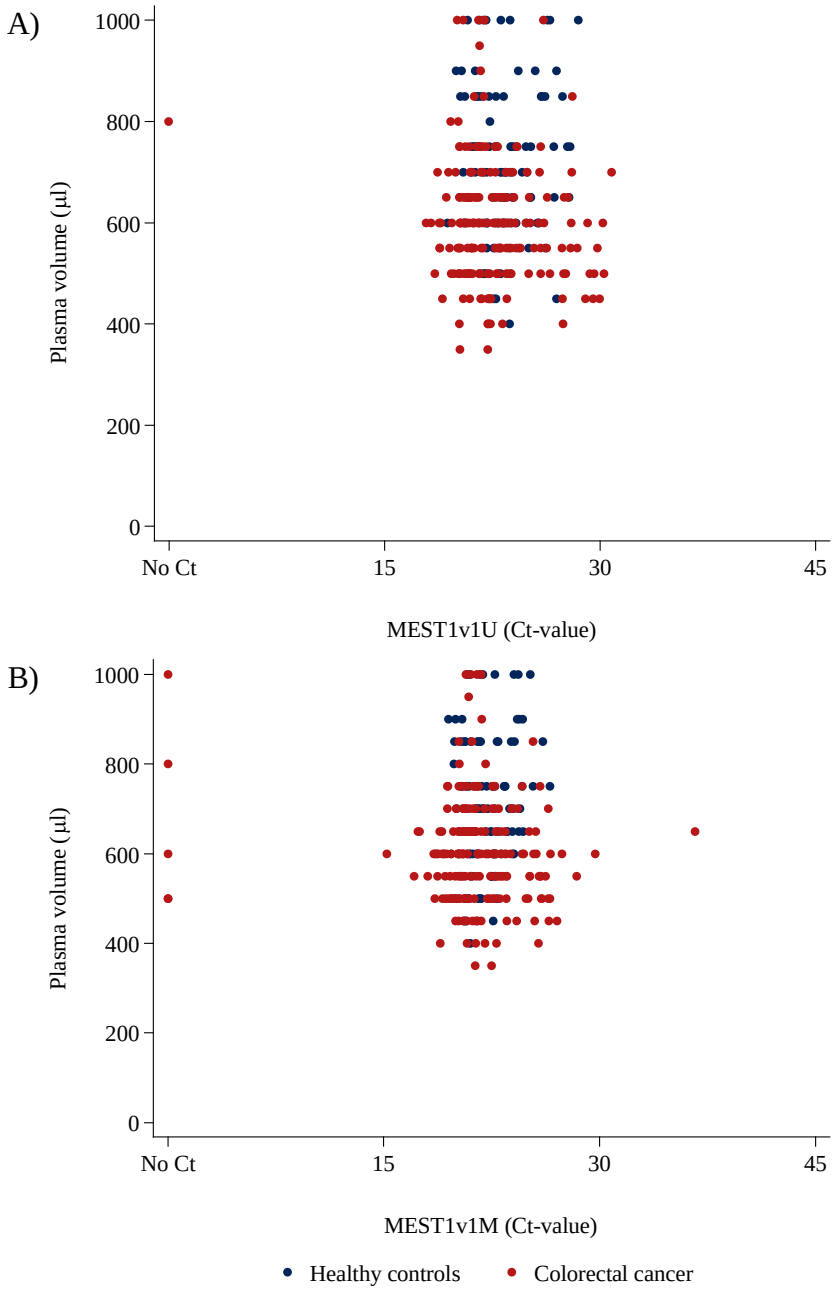
	CT value: Colorectal cancer patients (N=193)								*CT value: Healthy controls (N=102)*							
	No Ct		*0-25*		*25-30*		*>30*		*No CT*		*0-25*		*25-30*		*>30*	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
ALX4	138	71.5	8	4.1	20	10.4	27	14.0	101	99.0	0	0.0	0	0.0	1	1.0
APC	112	58.0	25	13.0	45	23.3	11	5.7	69	67.6	5	4.9	23	22.5	5	4.9
BMP3	138	71.5	18	9.3	21	10.9	16	8.3	91	89.2	1	1.0	3	2.9	7	6.9
BNC1	170	88.1	4	2.1	8	4.1	11	5.7	89	87.3	0	0.0	2	2.0	11	10.8
BRCA1	144	74.6	7	3.6	10	5.2	32	16.6	80	78.4	3	2.9	7	6.9	12	11.8
CDKN2A	175	90.7	7	3.6	3	1.6	8	4.1	98	96.1	0	0.0	2	2.0	2	2.0
HIC1	182	94.3	1	0.5	7	3.6	3	1.6	101	99.0	0	0.0	0	0.0	1	1.0
HLTF	171	88.6	8	4.1	9	4.7	5	2.6	98	96.1	1	1.0	1	1.0	2	2.0
MGMT	182	94.3	3	1.6	5	2.6	3	1.6	101	99.0	0	0.0	0	0.0	1	1.0
MLH1	106	54.9	20	10.4	46	23.8	21	10.9	58	56.9	2	2.0	20	19.6	22	21.6
NDRG4	175	90.7	1	0.5	1	0.5	16	8.3	102	100	0	0.0	0	0.0	0	0.0
NPTX2	58	30.1	30	15.5	61	31.6	44	22.8	42	41.2	5	4.9	25	24.5	30	29.4
NEUROG1	153	79.3	10	5.2	19	9.8	11	5.7	82	80.4	0	0.0	6	5.9	14	13.7
OSMR	171	88.6	9	4.7	10	5.2	3	1.6	95	93.1	0	0.0	2	2.0	5	4.9
PHACTR3	165	85.5	4	2.1	12	6.2	12	6.2	96	94.1	0	0.0	1	1.0	5	4.9
PPENK	173	89.6	3	1.6	3	1.6	14	7.3	98	96.1	0	0.0	0	0.0	4	3.9
RARB	144	74.6	8	4.1	34	17.6	7	3.6	31	30.4	2	2.0	23	22.5	46	45.1
RASSF1A	171	88.6	10	5.2	8	4.1	4	2.1	86	84.3	2	2.0	5	4.9	9	8.8
SDC2	146	75.6	12	6.2	26	13.5	9	4.7	96	94.1	0	0.0	0	0.0	6	5.9
SEPT9	146	75.6	8	4.1	21	10.9	18	9.3	97	95.1	0	0.0	0	0.0	5	4.9
SFRP1	151	78.2	12	6.2	16	8.3	14	7.3	95	93.1	0	0.0	4	3.9	3	2.9
SFRP2	154	79.8	26	13.5	7	3.6	6	3.1	84	82.4	5	4.9	8	7.8	5	4.9
SPG20	163	84.5	4	2.1	12	6.2	14	7.3	90	88.2	0	0.0	0	0.0	12	11.8
SST	135	69.9	31	16.1	24	12.4	3	1.6	70	68.6	5	4.9	23	22.5	4	3.9
TAC1	91	47.2	4	2.1	30	15.5	68	35.2	54	52.9	0	0.0	3	2.9	45	44.1
THBD	174	90.2	4	2.1	13	6.7	2	1.0	101	99.0	0	0.0	0	0.0	1	1.0
TFPI2	179	92.7	8	4.1	6	3.1	0	0.0	100	98.0	0	0.0	0	0.0	2	2.0
VIM	159	82.4	6	3.1	7	3.6	21	10.9	90	88.2	0	0.0	1	1.0	11	10.8
WIF1	174	90.2	0	0.0	4	2.1	15	7.8	98	96.1	0	0.0	0	0.0	4	3.9
WNT5A	181	93.8	1	0.5	4	2.1	7	3.6	97	95.1	0	0.0	0	0.0	5	4.9

Note. The number (N) and percentages (%) of colorectal cancer patients and healthy controls according to cycle threshold value (CT) in the polymerase chain reaction. "No CT" refers to no signal.

Supplementary figures

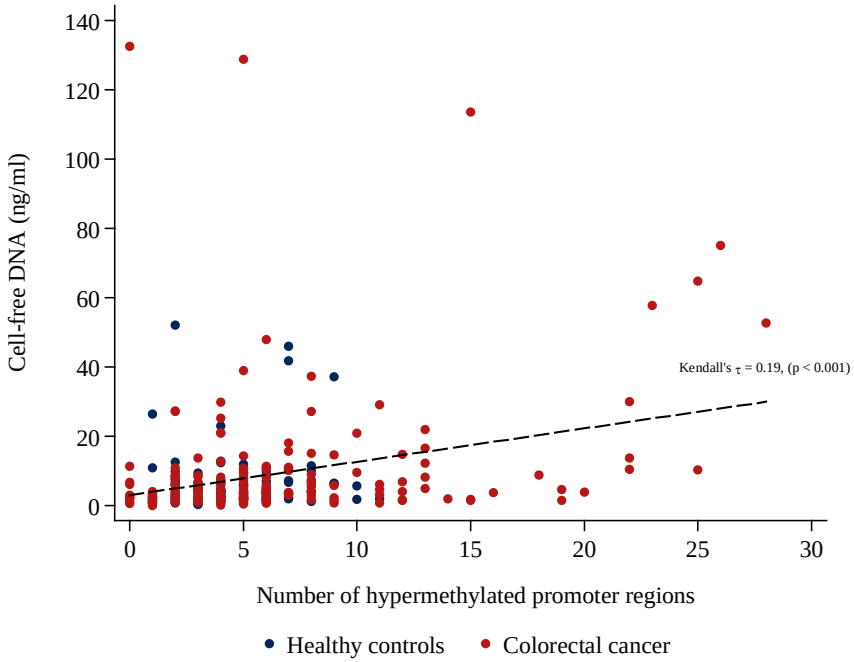


Supplementary figure 1 Hypermethylation analysis of cell-free plasma derived DNA
Schematic overview of the bisulphite treatment protocol described by Pedersen et al. (2012)



Supplementary figure 2 Association between cycle threshold and the reference gene

A) Sample volume according to cycle threshold (Ct) value of unmethylated *MEST1*. B) Sample volume according to Ct value of methylated *MEST1*. Five patients did not have amplification of the reference gene (noCT). These patients were excluded from further analysis



Supplementary figure 3 Cell-free DNA and hypermethylated DNA

The concentration of cell-free DNA according to the number of hypermethylated promoter regions measured in plasma at baseline. Healthy controls are marked in blue and colorectal cancer patients with red. The dashed line represents the correlation between the values. Kendall's rank coefficient (τ) is provided as a measure of the association.

Papers

I. Hypermethylated DNA, a Diagnostic and Prognostic Marker for Colorectal Cancer – a systematic review.
S. L. Rasmussen, H. B. Krarup, K. G. Sunesen, I. S. Pedersen, P. H. Madsen, O. Thorlacius-Ussing.
Colorectal Disease. 2016 Jun;18(6):549-61.
DOI: 10.1111/codi.13336

II. Hypermethylated DNA, a Circulating Biomarker for Colorectal Cancer Detection.
S. L. Rasmussen, H. B. Krarup, K. G. Sunesen, M. B. Johansen, M. T. Stender, I. S. Pedersen, P. H. Madsen, O. Thorlacius-Ussing.
PLoS One. 2017 Jul 10;12(7):e0180809
DOI: 10.1371/journal.pone.0180809

III. The Prognostic Utility of Circulating Hypermethylated DNA in Colorectal Cancer.
S. L. Rasmussen, H. B. Krarup, K. G. Sunesen, M. B. Johansen, M. T. Stender, I. S. Pedersen, P. H. Madsen, O. Thorlacius-Ussing
Submitted for publication in *Oncotarget*.

ISSN (online): 2246-1302
ISBN (online): 978-87-7210-100-2

AALBORG UNIVERSITY PRESS